Universidad Autónoma de Madrid

Facultad de Ciencias

Departamento de Biología Molecular

ICETh1 & ICETh2, two mobile genetic elements coordinated in *Thermus thermophilus* transjugation

TESIS DOCTORAL

Ignacio Baquedano Mozos

Madrid, 2019

Universidad Autónoma de Madrid

Facultad de Ciencias

Departamento de Biología Molecular

ICETh1 & ICETh2, two mobile genetic elements coordinated in *Thermus thermophilus* transjugation

Memoria presentada por Ignacio Baquedano Mozos, licenciado en Biología, para optar al título de Doctor en Biociencias Moleculares por la Universidad Autónoma de Madrid

Directores:

José Berenguer Carlos

Mario Mencía Caballero

Tutor:

José Berenguer Carlos

Centro de Biología Molecular Severo Ochoa

Madrid, 2019

A mis padres

Acknowledgements

Han sido cuatro años, a lo largo de los cuales, he aprendido mucho, y no solo científicamente hablando. Han sido muchas personas las que han pasado por mi vida en este periodo y me gustaría destacar a algunas en concreto.

Empezando por el lugar donde se ha desarrollado la tesis tengo que destacar a los que están y a los que ya emprendieron otro camino. Gracias a mis compañeros de laboratorio por la ayuda y por compartir experiencias, cafés y almuerzos y por su trabajo, el cual ha conseguido que siempre haya financiación para realizar los diferentes experimentos. Todos han sido relevantes, pero sin duda, tengo que destacar a Sandra, a Carlos, a Jorge, a Marcos, a Diana y a María Luisa y a algunas adquisiciones más recientes que han sido de ayuda como Elena o Dione. También destacar el compañerismo de laboratorios cercanos. De entre ellos destaco a Peter, Zoran, Tamara y Patri, pero sin olvidarme de los demás.

Mis experiencias previas de laboratorio también han sido de gran importancia. Destaco mi periodo en el CNB, en el cual, básicamente aprendí a trabajar en un laboratorio de microbiología. Aquí tuve grandes compañeros y buenos amigos a su vez, pero tengo que destacar el papel de mi tutora Silvia, ya que fue ella quien tuvo la paciencia de enseñarme.

Destacar también el papel de mis directores de tesis, Pepe y Mario, los cuales me han guiado con muy buen criterio y apoyo a lo largo de este recorrido. También agradecer a los colaboradores de esta tesis, Alba y Sandra González su trabajo y esfuerzo. Y no me puedo olvidar del servicio de genómica y secuenciación masiva donde tengo que agradecer en especial el trabajo de Laura.

Tengo que agradecer también la financiación aportada por el ministerio de ciencia (antiguamente MINECO). Además, con dicha financiación pude desplazarme a Sherbrooke en Canadá para aprender sin duda de una de las personas que más sabe de ICEs, de Vincent Burrus, que me acogió en su grupo de trabajo. En estos tres meses aprendí, disfruté mucho y conocí a grandes personas entre las cuales tengo que destacar a Kevin.

Pero no todo es ciencia. Fuera de este ámbito tengo que agradecer a mucha gente por muchos buenos momentos. Sin duda a mis “boques”, Dani, Edu y Álvaro, ese tándem inseparable. Y, sin olvidarme, por supuesto de Noe y Tania sin las cuales no se entendería este tándem. Agradecer también a mis amigos del barrio, es decir, a mis amigos de toda la vida (Richi, Carlos, Fernando, María, Alex, Irene y Borja), que sin duda alguna también han estado ahí, compartiendo momentos. Y a mis compañeros de facultad Gonzalo, Sara, Lorea y Alex que, aunque por temas de distancia no es fácil reunirnos, son un apoyo. Y no puedo olvidarme del mundo del tenis, donde puedo encontrar siempre a dos fenómenos que comparten cenas, viajes, partidos y experiencias. Gracias Rubén, gracias Miguel.

Por último, dar gracias a todos aquellos que están en mi día a día, en todos los momentos, en los buenos y en los malos. Gracias a Dios, gracias a mi familia, sobre todo a mis padres que han estado, están y estarán siempre brindándome un apoyo incondicional sin importar el esfuerzo, o el tiempo que lleve, pero también a mi hermano y a Sabrina y finalmente mi abuela.

Finalmente, no puedo olvidarme de alguien que, aunque solo ha estado conmigo durante el último cuarto de la tesis me ha soportado sin queja alguna y que sin duda me ha hecho ver las cosas de otra manera y sonreír. Gracias Anabel.

Gracias a todos.

Table of contents

Table of contents

Acknowledgements VI

Table of contents XI

Abbreviations XVII

Summary XXII

Resumen XXVII

Chapter 1: Introduction 3

1.1 Horizontal gene transfer 3

1.1.1 Mechanisms of HGT 4

1.1.1.1 Transformation 4

1.1.1.2 Transduction and other DNA-protected HGT mechanisms 5

1.1.1.3 Conjugation 6

1.1.1.4 Unconventional conjugation models 8

1.1.2 Integrative and Conjugative Elements 11

1.1.3 HGT barriers 15

1.2 The *Thermus* genus 16

1.2.1 HGT in *Thermus* spp. 17

1.2.1.1 Transformation 18

1.2.1.2 Transjugation 19

1.2.1.3 Transduction 21

1.2.1.4 Vesicle-protected HGT 21

1.2.2 Barriers against HGT in *Thermus* 22

Chapter 2: Scope and objectives 26

Chapter 3: Materials and Methods 30

3.1 Materials 30

3.2 Microbiological methods 35

3.2.1 Bacterial growth and storage conditions 35

3.2.2 Bacterial transformation 35

3.2.3 Toxicity assays 36

3.2.4 Transjugation assays 37

3.2.5 Stress induction assays by UV light 37

3.3 Molecular methods 37
3.3.1 DNA manipulation 37
3.3.2 RNA manipulation 38
3.3.3 Mutants generation 39
3.4 Nucleic acids amplification 40
3.4.1 Conventional PCR 40
3.4.2 qPCR (in collaboration with genomics and NGS facility at CBMSO) 40
3.4.3 RT-PCR (in collaboration with genomics and NGS facility at CBMSO) 42
3.5 Processing of samples for microscopic visualization 43
3.5.1 In *vivo* fluorescent staining 43
3.5.2 Sorting and confocal microscopy 43
3.6 Transjugants analysis 43
3.6.1 Assemblies and annotation sets 44
3.6.2 Alignment and pre-process of alignment files 44
3.6.3 Variant calling 44
3.6.4 Identification of potential variants, filter, number and annotation of discovered variants 44
3.6.5 Results analysis 45
3.6.6 Circos representation of variants against HB27 or HB8 genome and plasmid sequence 45
3.6.7 Resistance genes location: *kat* and *hyg* genes 45
3.6.8 Strain-specific genes of donor strain in transjugant samples 45
3.6.9 Analysis of unmapped reads vs HB8 genome reference 46
3.7 Bioinformatic toolbox 46
Chapter 4: Identification and bioinformatics analysis of ICEThs of *T. thermophilus* 50
4.1 ICETh1 analysis 50
4.2 ICETh2 analysis 52
4.3 ICETh3 analysis 54
4.4 Comparison of putative ICEThs 56
Chapter 5: ICEThs excision, transfer and copy number 60
5.1 Excision of ICETh1 and ICETh2 from the chromosome 60
5.2. ICETh1 and ICETh2 transfer to recipient cells 65
Chapter 6: Excision/integration and transfer modules 70
6.1 Excision/integration modules 70
6.1.1. Mobilization of ICETh1 depends on ICETh2 71

6.1.2 *int1* and *int2* are transcribed at low levels 74
6.1.3. Excision of ICETh1 affects to its intracellular mobilization 75
6.2 Transfer module 76
6.2.1 Transfer module is cotranscribed 76
6.2.2 Mutants and their effects 77
6.2.3 Absence of a restriction site for transfer 78
Chapter 7: replication and accessory modules 83
7.1 Replication module 83
7.1.1 ICEThs excision is promoted by stress 83
7.1.2 Putative role of PrimPol in replication 85
7.1.3 Analysis of the role of the TOPRIM domain homologue *TTC0657* 87
7.1.4 Role of *TTC0658* 88
7.2 Regulation module 90
7.3 Maintenance module 91
Chapter 8: Transjugation-generated mosaicity 97
Chapter 9: Discussion 110
Conclusions 130
Conclusiones 135
Bibliography 140
Online resources 158
Annex I 162
Annex II 170

Abbreviations

Abbreviations

Ab^R	Antibiotic resistance cassette
Ago	Argonaute protein
Amp	Ampicilin
Amp^R	Ampicilin resistant
attB	Attachment site in the chromosome
attI	Attachment site in the circularized ICE
attL	Left attachment site of the integrated ICE
attR	Right attachment site of the integrated ICE
Bleo	Bleomycin
Cm	Chloramphenicol
Cm^R	Chloramphenicol resistant
Cq	Cycle quantification
DCT	Distributive Conjugal Transfer
eDNA	Extracellular DNA
EVs	Extracellular vesicles
GFP	Green Fluorescent Protein
GTA	Gene Transfer Agents
HGT	Horizontal Gene Transfer
HTH	Helix-Turn-Helix DNA binding motif
Hyg	Hygromycin
hyg	Hygromycin resistance cassette
Hyg^R	Hygromycin resistant

ICE	Integrative and Conjugative Element
ICETh	Integrative and Conjugative Element of *Thermus thermophilus*
IM	Inner Membrane
IME	Integrative and Mobilizable Elements
kat	Kanamycin resistance cassette
Km	Kanamycin
Km^R	Kanamycin resistant
LB	Luria Bertani medium
MGE	Mobile Genetic Elements
MV	Membrane Vesicles
NCA	Natural Competence Apparatus
OM	Outer Membrane
ORF	Open Reading Frame
oriT	Origin of transfer
pAgo	Prokaryotic Argonaute
PG	Peptidoglycan
RCR	Rolling Circle Replication
RDF	Recombination Directionality Factor
RM	Restriction-Modification
RT-PCR	Real time PCR
SD	Standard Deviation
sGFP	Superfolder Green Fluorescent Protein
sIFP	Thermostable version of Green Fluorescent Protein

SL	S-Layer
SOC	Super Optimal Broth with Catabolite Repression
T4P	Type IV Pilus
T4SS	Type IV Secretion System
TA	Toxin-Antitoxin
TB	*Thermus* broth
tDNA	DNA to be transferred
TtAgo	*Thermus thermophilus* Argonaute
UV	Ultraviolet light
WGS	Whole Genome Sequence
wt	Wild type

Summary

Summary

Horizontal Gene Transfer (HGT) is considered one of the most important sources of bacterial evolution. In the thermophilic bacterium *Thermus thermophilus*, the transjugation mechanism has been described as a highly efficient HGT system. This mechanism consists in the transfer of DNA from a donor to a recipient cell, being necessary the machinery for DNA donation in the donor and the transformation machinery in the recipient cell.

In the strain HB27 of this organism, two small Integrative and Conjugative Elements (ICEs), have been discovered and we show how they coordinate their activities in the transjugation process. ICETh1 encodes the machinery necessary for DNA donation, where the translocase TdtA is essential, as it is probably the most important protein in this process. However, accessory proteins such as the nuclease NurA and the restrictase Tth111II are also relevant for this event, likely processing DNA prior to its transfer mediated by TdtA. Any locus in the genome can be transferred to a recipient cell; however, the ICEThs show a higher transfer rate.

We have shown that ICETh1 is not capable by itself to excise or integrate in the chromosome. This process is dependent on a specialized excision/integration module encoded in ICETh2, a second mobile element that can excise and integrate both ICEThs in their respective sites despite being catalyzed by a single enzyme

These ICEThs seem to exhibit a higher excision rate and apparently replicate under stress conditions produced by UV or in the absence of the primase/polymerase PrimPol encoded by ICETh2. Autonomous replication of both ICEThs, even when it is not clear, could be driven by a TOPRIM-domain homologue protein.

A Toxin-Antitoxin (TA) system could assure the presence of ICETh1 in the cell population via post-segregational killing. Additionally, the possibility exists that the TOPRIM-domain homologue, encoded in ICETh2, could be required for some cellular function promoting somehow the maintenance of ICETh1 and, to a lesser extent of ICETh2.

Furthermore, in this work it is proposed a retro-transfer model for DNA transjugation in which DNA fragments, apart from the ICEThs themselves, can be transferred to a

recipient cell in which they integrate. Then, with the intervention of ICETh1 transjugation machinery, DNA fragments from the recipient cell could be transferred back in the opposite direction to the original donor, followed by integration, generating mosaicism in the progeny.

Resumen

Resumen

La transferencia horizontal de genes (HGT) está considerada como uno de los mecanismos más relevantes en la evolución bacteriana. En la bacteria termófila *Thermus thermophilus*, el mecanismo de transjugación ha sido descrito como un sistema altamente eficiente de transferencia horizontal de genes. Este mecanismo consiste en la transferencia de ADN de una célula donadora a una célula receptora, siendo necesaria la maquinaria de transferencia de DNA en la donadora y la maquinaria de transformación en la receptora.

En la cepa HB27 de este organismo, dos elementos integrativos conjugativos (ICEs) de pequeño tamaño han sido descubiertos. En el presente trabajo se describe como ambos coordinan su actividad en el proceso de transjugación. ICETh1 codifica la maquinaria necesaria para el proceso de donación del DNA, en el cual la translocasa TdtA es esencial, siendo probablemente la proteína más importante en dicho proceso. Sin embargo, una serie de proteínas como la nucleasa NurA y la restrictasa Tth111II son también relevantes en este proceso, en el cual, probablemente se encarguen del procesamiento del ADN antes de su transferencia mediada por TdtA. En este proceso de transjugación, cualquier locus en el genoma puedes ser transferido a una célula receptora; sin embargo, los ICEThs exhiben una mayor tasa de transferencia.

Se ha demostrado en este trabajo que ICETh1 no es capaz de escindirse o integrarse en el cromosoma bacteriano por sí mismo. Este proceso es llevado a cabo por un módulo especializado de escisión/integración codificado en ICETh2, un segundo elemento móvil. Este módulo cataliza, con una única enzima, tanto la escisión como la integración de ICETh1 e ICETh2 en sus sitios correspondientes de recombinación.

Ambos ICEThs parecen mostrar una mayor tasa de escisión o incluso replicar bajo condiciones de estrés producidas por UV o por la ausencia de la primasa/polimerasa PrimPol. En caso de producirse la replicación autónoma de ambos elementos, esta podría estar mediada por una proteína con homología al dominio TOPRIM.

Un sistema Toxina-Antitoxina (TA) podría asegurar la presencia de ICETh1 en la población a través de la destrucción de aquellas bacterias que pierdan el elemento.

Adicionalmente, existe la posibilidad de que el homólogo del dominio TOPRIM, codificado en ICETh2, pueda tener alguna función celular que promueva de alguna manera el mantenimiento en la población de ICETh1, y en menor medida de ICETh2.

Por último, en esta tesis se propone un modelo de retrotransferencia para la transjugación en el cual, los fragmentos de ADN, además de los ICEThs, serían transferidos a una célula receptora en la cual se integrarán. A continuación, con la intervención de la maquinaria de transjugación codificada en ICETh1, otros fragmentos de ADN de la célula receptora serían transferidos en dirección contraria hacia el donador original, integrándose en este y generando, consecuentemente, mosaicismo en la progenie.

Chapter 1

Introduction

Chapter 1: Introduction

1.1 Horizontal gene transfer

Bacteria display a vast range of genome sizes, cellular structures, metabolic properties and lifestyles. Even related taxonomic groups show unique physiological characteristics. Several mechanism have been shown to be responsible for these differences, including point mutations (Amábile-Cuevas and Chicurel, 1993). However, it is difficult to assume that this mutation mechanism by itself drives the ability of bacteria to exploit new environments. Instead, there is a growing evidence that Horizontal Gene Transfer (HGT) plays an important role in the evolution, diversification and speciation of bacteria (Ochman *et al.*, 2000). Traditionally, three major mechanisms for HGT have been studied in bacteria: acquisition of exogenous DNA (eDNA) or transformation, transfer of DNA from a donor to a recipient cell or conjugation and DNA transfer mediated by bacteriophages or transduction. However, alternative mechanisms such as membrane vesicles or gene transfer agents (GTA) among others have been reported in prokaryotes (Fig 1.1).

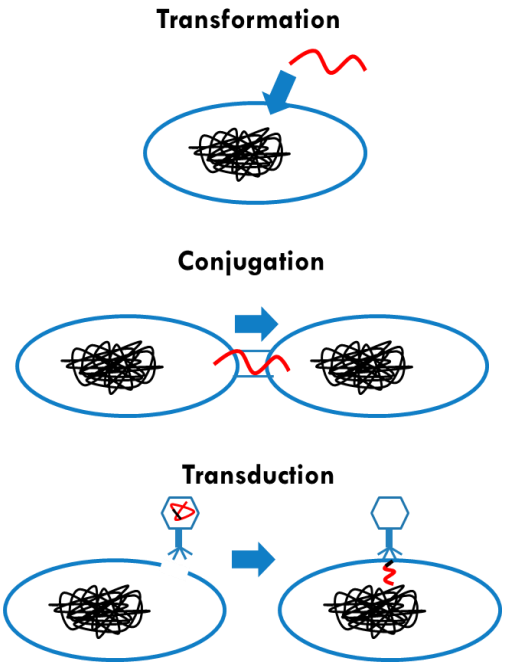


Figure 1.1: Main mechanisms of horizontal gene transfer. Transformation: uptake of eDNA; Conjugation: transfer of DNA fragments between cells that are in contact; Transduction: transfer of chromosomal fragments mediated by bacteriophages. Chromosomal DNA is depicted as a tangled line and DNA to be transferred in red.

1.1.1 Mechanisms of HGT

1.1.1.1 Transformation

This mechanism was discovered in *Streptococcus pneumoniae* in 1928 (Griffith, 1928). Naturally transformable bacteria present a physiological state known as "competence" which just a few bacterial and archaeal species can acquire (Mell *et al.*, 2012; Johnston *et al.*, 2014). During this state, which can be achieved by different conditions depending on the bacterial species (Aas *et al.*, 2002; Berka *et al.*, 2002; Claverys *et al.*, 2006; Veening and Blokesch, 2017) they are capable of uptaking eDNA from the environment (Berka *et al.*, 2002; Ogura *et al.*, 2002). The mechanism for natural transformation is well conserved among Gram-positives (G+) (*S. pneumoniae* and *Bacillus subtilis* for example) and Gram-negatives (G-) (e.g. *Neisseria gonorrhoeae*, *Vibrio cholera* and *Haemophilus influenzae*) laboratory models (Chen and Dubnau, 2004; Claverys *et al.*, 2009; Burton and Dubnau, 2010; Johnston *et al.*, 2014; Cabezón *et al.*, 2015; Ilangovan *et al.*, 2015; Veening and Blokesch, 2017). eDNA uptake involves highly conserved proteins (Johnston *et al.*, 2014) except, among others, for *Helicobacter pylori*, in which this DNA uptake is performed by a conjugation-like system (Smeets and Kusters, 2002).

In general, double-stranded eDNA is pulled across the outer membrane (OM) (in G- bacteria), then across the peptidoglycan (thicker and denser in G+) and it is finally internalized to the cytoplasm across the inner membrane (IM) as single-stranded DNA (ssDNA).

Specifically, in the G- bacteria *Neisseria gonorrhoeae*, during the assembly and disassembly of a type IV pilus (T4P), a pseudopilus is formed (integrated by pseudopilins multimeres) that reaches to the OM where a pore is present, formed by the secretin PilQ (Chen and Dubnau, 2004). This pore is large enough to accommodate double-stranded DNA (dsDNA) (Collins *et al.*, 2001; Assalkhou *et al.*, 2007; Burkhardt *et al.*, 2011). When T4P retracts, the attached-dsDNA crosses the OM through the PilQ secretin (Laurenceau *et al.*, 2013; Salzer *et al.*, 2014, 2016; Leong *et al.*, 2017). In the periplasm, incoming DNA is bound to the ComEA protein (Inamine and Dubnau, 1995; Bergé *et al.*, 2002; Seitz *et al.*, 2014) and it is translocated across the IM through a pore formed by the well conserved ComEC protein (Bergé *et al.*, 2002; Draskovic and Dubnau, 2005). This protein acts as a single-stranded nuclease as well (Chen and Dubnau, 2004; Claverys

et al., 2009; Burton and Dubnau, 2010) degrading one of the DNA stands while the other strand is translocated to the cytoplasm. Energy for this translocation is provided by the ComFA ATPase (Londoño-Vallejo and Dubnau, 1994; Takeno *et al.*, 2011). Finally, after translocation of the ssDNA, DprA and RecA proteins bind to the ssDNA and in the case of non plasmidic DNA, starts the recombination process with the genome (Mortier-Barrière *et al.*, 2007; Yadav *et al.*, 2013, 2014; Duffin and Barber, 2016) (Fig. 1.2).

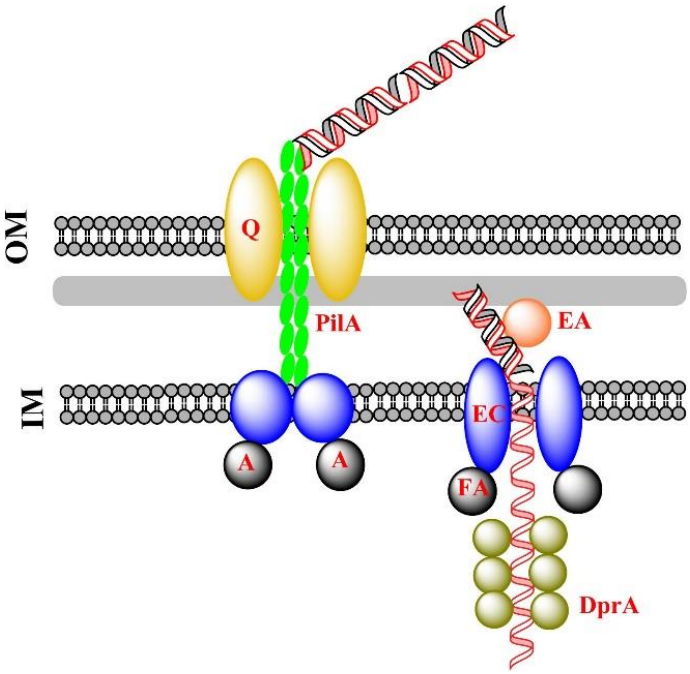


Figure 1.2: Transformation machinery in G-. dsDNA is internalized by the extension and retraction of pseudopili formed by pseudopilin multimeres (PilA) and pulled through PilQ pore (G- only). ComEA binds dsDNA and transfer one strand (degrading the other) with the assistance of the ATPase ComFA. ssDNA in the cytoplasm is bound to DprA (Sun, 2018).

1.1.1.2 Transduction and other DNA-protected HGT mechanisms

Bacteriophages play an important role in shaping the bacterial microbiome in any environment (Von Wintersdorff *et al.*, 2016). In the transduction process, DNA transfer is mediated by bacteriophages that encapsidate DNA fragments from the host accidentally during the infection cycle and then, after a new infection, the fragments can be inserted in a new host. Through this mechanism, bacteriophages can transfer genes advantageous for their host and consequently, promote their own survival and dissemination (Modi *et al.*, 2013).

In the case of Gene transfer agents (GTAs), first described in *Rhodobacter capsulatus*, short fragments of dsDNA from the host genome of this α-Proteobacteria are packaged randomly in host-cell produced particles that resemble bacteriophages structures (Marrs, 1974; Humphrey *et al.*, 1997). GTAs are unable of self-propagation as the

amount of DNA packaged by the GTAs is insufficient to encode all of their protein components (Lang and Beatty, 2000, 2001, 2007). These GTA particles are released through cell lysis (Hynes *et al.*, 2012; Westbye *et al.*, 2013) and it has been proposed that for cell entry they require proteins involved in natural transformation (Brimacombe *et al.*, 2015).

Phages play also a relevant role in *Staphylococcus aureus*, being transduction its major source of genetic variation (Viana *et al.*, 2015) and its main horizontal gene transfer mechanism (Penadés and Christie, 2015). An example of the role of staphylococcal phages in spreading virulence is observed with the *S. aureus* pathogenicity islands (SaPIs) whose transfer relies on the phage machinery upon infection (reviewed in Penadés and Christie, 2015).

Extracellular Vesicles (EVs) or Membrane Vesicles (MVs) are membrane-enclosed structures (produced in prokaryotes) produced by bacteria, archaea and eukaryotes during the growth (Turnbull *et al.*, 2016) in which eDNA and/or cytoplasmic enzymes can be encapsulated playing a great diversity of roles, such as cell communication, immune system evasion, and stress response, among others (Schwechheimer and Kuehn, 2015). This mechanism is quite common in G- bacteria, like in *Pseudomonas aeruginosa*. How these vesicles enter recipient cells is not known, but its fusion with the OM could release the content in the periplasmic space and an unknown mechanism would internalize it, as proposed by Fulsundar *et al.*, 2014 likely with the help of the NCA (natural competence apparatus).

1.1.1.3 Conjugation

This mechanism, discovered by Lederberg and Tatum in 1946, relies on cell-to-cell contacts. In this context, DNA is pushed out of a donor cell and transported into a recipient cell. The transfer of DNA across the membrane of the donor bacterium relies on a large membrane-associated protein complex which belongs to the type IV secretion system (T4SS) (Goessweiner-Mohr *et al.*, 2013; Cabezón *et al.*, 2015; Ilangovan *et al.*, 2015).

This mechanism has initially been described for model *Proteobacteria*. It is mediated by two functional modules encoded in mobile genetic elements (MGE): either in self-replicative conjugative plasmids, or in mobilizable plasmids assisted by proteins encoded in the chromosome; or in Integrative and Conjugative Elements (ICEs) inserted in the chromosome (see 1.1.2) (Guglielmini *et al.,* 2011; Johnson and Grossman, 2015). Of the two modules, one codes for DNA processing and replication (Dtr) and the other for the mating pair formation (Mpf) in which the T4SS is encoded (Smillie *et al.*, 2010).

Proteins encoded in the Dtr module perform the DNA processing, transfer and coupling to the T4SS. Then, the transfer of the DNA between two cells that are in contact by the conjugative pilus can start. The pilus is formed by pilins (Anthony *et al.*, 1999) whose depolimerization in the base brings closer both mates. The Dtr module involves a *cis*-acting DNA sequence named the origin of transfer (*oriT*) and a relaxase that recognizes this *oriT* and then, produce a cut on one of the strands, remaining covalently bound to its 5' end. An auxiliary transfer factor binds to *oriT* to ensure the specificity of the cleavage. The complex formed by the relaxase and auxiliary transfer factors bound to *oriT* is known as the relaxosome complex (Furste *et al.*, 1989). Simultaneously, the complementary strand is replicated via rolling circle replication (RCR). This relaxosome complex is recruited to the channel of the T4SS, where it docks to the coupling protein whose ATPase activity allows the translocation of the relaxase bound to the tDNA (DNA to be transferred) into a passive recipient cell (Llosa *et al.*, 2002; Koraimann and Wagner, 2014; Cabezón *et al.*, 2015; Christie, 2016). Finally, once the relaxase-tDNA complex has been translocated, the relaxase re-circularizes the transferred strand through DNA ligation (Draper *et al.*, 2005). The process is depicted in Figure 1.3.

In case that the machinery necessary for conjugation was encoded in the chromosome giving rise to high frequency of recombination (Hfr) cells (Tatum and Lederberg, 1947; Curtiss and Renshaw, 1969), part of the chromosome is transferred and generally interrupted by random breakage, impeding further transfer (Cavalli-Sforza, 1950).

The case of ICEs will be explained in section 1.1.2.

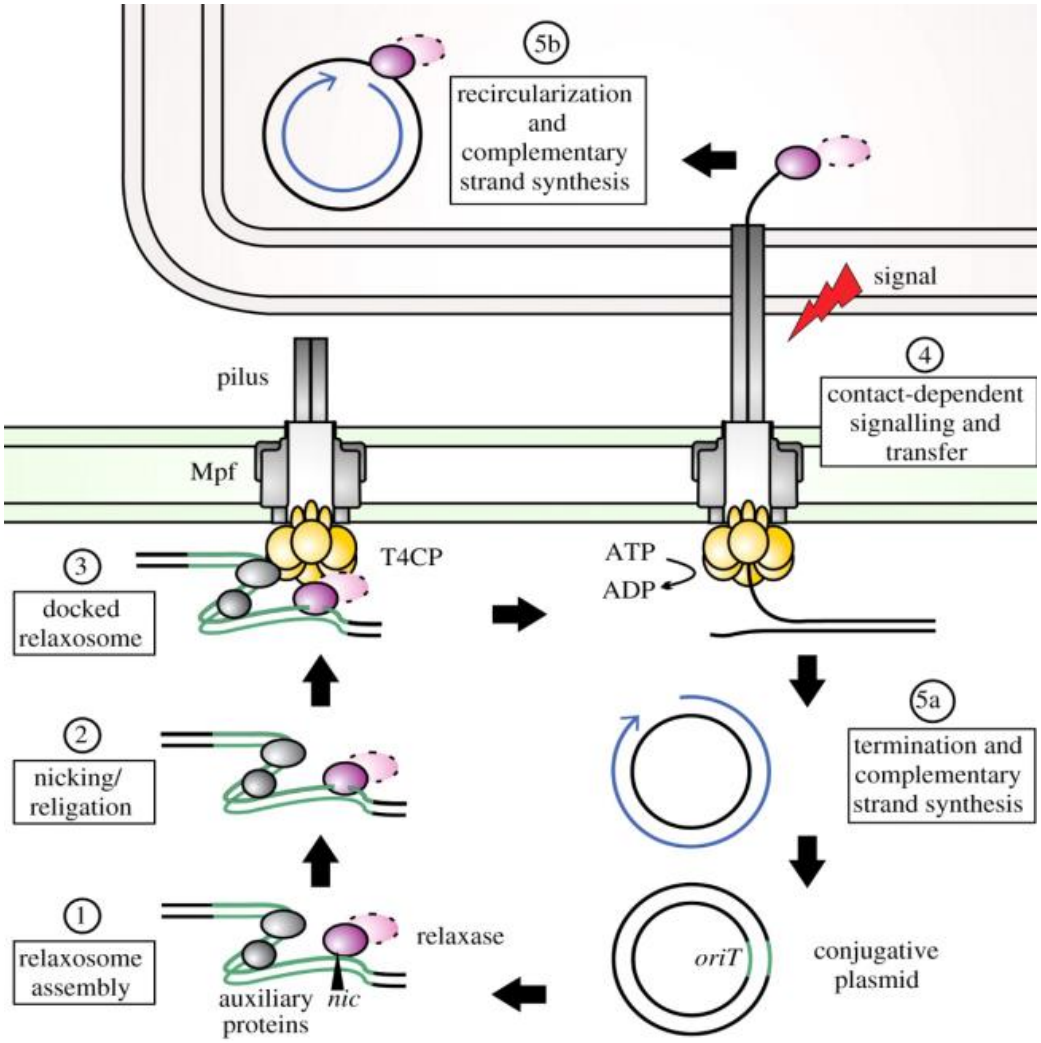


Figure 1.3: Conjugation mechanism: Relaxase (in purple) binds the *oriT* covalently with the assistance of auxiliary factors forming the relaxosome. Then it is attached to the coupling protein (in yellow) and transferred to the recipient cell through the channel created by the T4SS (in grey). (Hayes, 2001).

1.1.1.4 Unconventional conjugation models

Alternative models to this traditional conjugation mechanism have been reported in bacteria such as *Streptomyces*, *Mycobacteria* and *Mollicutes* (Vogelmann *et al.*, 2011; Derbyshire and Gray, 2014; Thoma and Muth, 2015; Citti *et al.*, 2018).

In *Streptomyces* spp., the small-sized conjugative plasmids encode a DNA translocase (TraB) which does not require any auxiliary proteins for primary transfer of the plasmid to a recipient cell (Thoma and Muth, 2012). This translocase, belonging to the FtsK family, forms hexameric structures on the membrane and through a C-terminal DNA

recognition motif, recognizes a specific cis-acting locus of transfer (*clt*) in the plasmid and, with lower affinity, in the chromosome. Experiments using restriction sites in the tDNA and structural features such as the size of the internal channel support that TraB transfers dsDNA to the recipient cells in an ATP dependent manner. Actually, TraB is the only protein needed for intermycelial transfer in *Streptomyces*. In addition to TraB, a few Spd proteins are also encoded in the conjugative plasmid of *Streptomyces*, which are required for further intramycelial transfer across the septa, allowing the rapid colonization of the whole mycelium (Thoma and Muth, 2015) (Fig. 1.4).

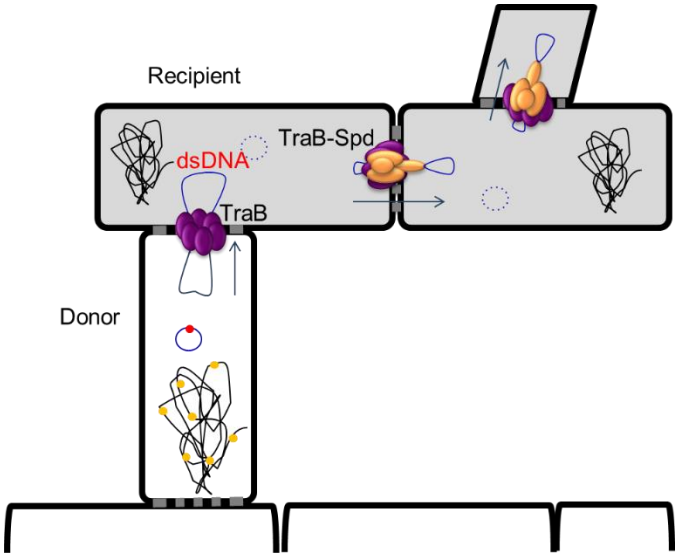


Figure 1.4: dsDNA transfer in *Streptomyces*: TraB recognizes and bind to *clt*. dsDNA is transferred to the recipient through the hexamer structure formed by TraB. Spd proteins collaborate in intramycelial transfer (Blesa and Berenguer, 2019).

In *Mycobacteria*, transfer of DNA between cells that are in contact occurs in the absence of known plasmids or ICEs. In this HGT model, named Distributive Conjugal Transfer (DCT), non-contiguous fragments of chromosomal DNA are transferred simultaneously, regardless of their chromosomal location. Consequently, the progeny shows a mosaic-like architecture derived from both parental strains, becoming some of them transconjugant donor cells (Wang *et al.*, 2005). These mosaic progeny are generated from a single conjugal event, which provides enormous capacity for rapid adaptation and evolution (Derbyshire and Gray, 2014). DCT capacity in *Mycobacteria* rely on the type IV secretion systems ESX-1 and ESX-4, being ESX-1 required in the donor for mating identity and ESX-4 essential in the recipient cell. Furthermore, donor cell contact is detected by recipient cells, resulting in induction of ESX-4, which enables the recipient to take up donor DNA during conjugation (Gray *et al.*, 2016; Gröschel *et al.*, 2016; Gray

and Derbyshire, 2018). Enzymes involved in fragmentation, recognition of origins of transfer are completely unknown in this model (Fig. 1.5).

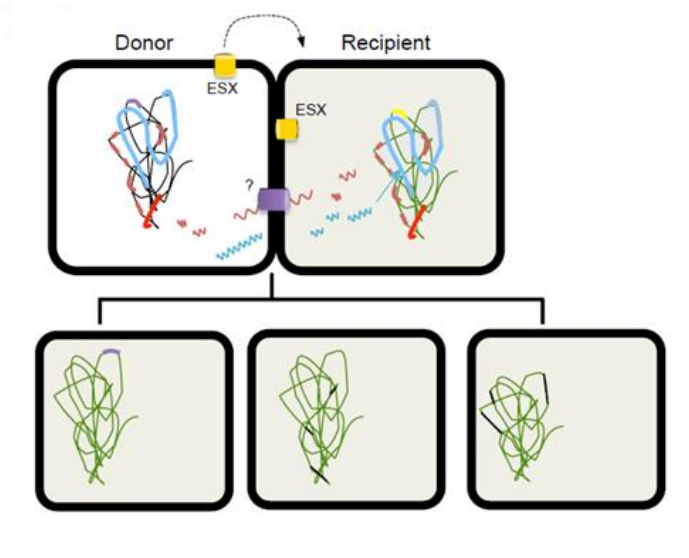


Figure 1.5. DCT model in *Mycobacteria*: Fragments of non-contiguous DNA are transferred from a donor to a recipient thanks to ESX1 and ESX4 secretion systems generating mosaicism in the progeny (Blesa and Berenguer, 2019).

Another unconventional conjugation mechanism occurs in *Mycoplasma*, specifically in *M. agalactiae* as model organism. In this case, the machinery associated to this process rely on an ICE (MICE). This element, besides typical genes encoded by ICE for their lifecycle (see 1.1.2), encodes a coupling TraG protein, a ssDNA binding protein and a TraE homologue, which likely participates in DNA transport (Citti *et al.*, 2018). During mating, transfer of large DNA fragments occur following two pathways. In one of them MICE is transferred from the ICE-positive cell (ICE+) to the ICE-negative cell (ICE-) while the transfer of the chromosomal DNA occurs in the opposite direction (Citti *et al.*, 2018) as seen in Figure 1.6A.

The mechanism involved in this process is unknown. However, a model has been proposed (Citti *et al.*, 2018) in which, despite no putative *oriT* sites have been identified, a yet unreported relaxase-like protein would mobilize one strand of the excised ICE to a secretion channel simpler than T4SS (as this secretion system is not present in *Mycoplasma*). Through this channel, a ssDNA-relaxase complex would be transferred to the recipient cell. In this process, cell-to-cell attachment is mediated by a surface lipoprotein instead of a T4P (Citti *et al.*, 2018) (Fig. 1.6B).

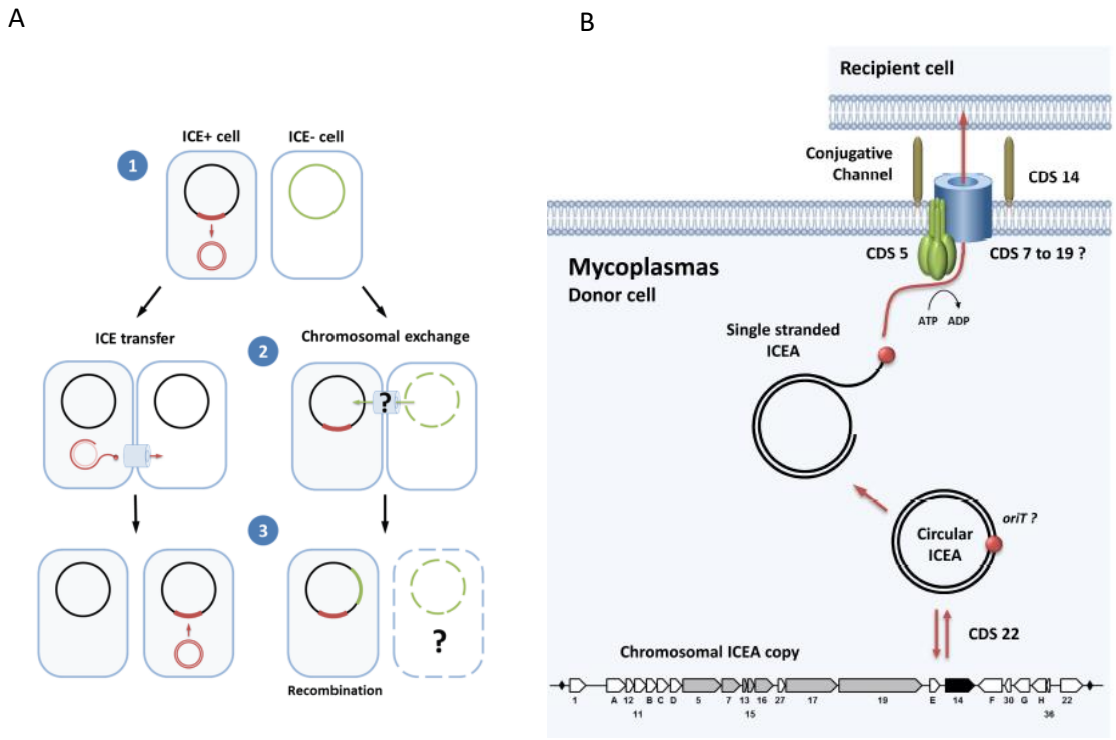


Figure 1.6. HGT in *Mycoplasma*: A. ICE is transferred from the ICE+ cell to the ICE- cell, while chromosomal fragments are transferred in the opposite direction. B. MICE encode the machinery essential for its transfer to the recipient cell (Citti *et al.*, 2018).

1.1.2 Integrative and Conjugative Elements

As mentioned in section 1.1.1.3, the conjugation machinery can be encoded either in conjugative plasmids or in ICEs (Guglielmini *et al.,* 2011; Johnson and Grossman, 2015). The essential difference between them relies in their respective modes of maintenance in a bacterial cell. While conjugative plasmids replicate autonomously for long-term persistence in cell populations, ICEs (formerly called conjugative transposons) integrate in the chromosome and replicate along with it (Wozniak and Waldor, 2010;Johnson and Grossman, 2015). The latest are integrated at specific sites in the chromosome (*attB*) of their host. This integration site corresponds usually, to the 3' end of tRNA or protein-encoding genes, which leads to the duplication of the integration sequence (*attL* and *attR*). ICEs are able to excise from the chromosome under specific circumstances generating an *attI* site on the circularized ICE and, consequently an *attB* site or "scar" in the chromosome. Like conjugative plasmids, the excised circular form of the ICEs serves as the substrate for transfer to the recipient cell mediated by the conjugative machinery.

Once in the recipient cell, ICEs integrate at their specific *attB* sites. Several ICEs have demonstrated their ability to replicate prior to transfer to a recipient cell (Grohmann, 2010; Lee *et al.*, 2010; Carraro *et al.*, 2015; Wright and Grossman, 2016). ICEs lifecycle is depicted in Figure 1.8.

ICEs are practically ubiquitous in prokaryote genomes, outnumbering conjugative plasmids (Ghinet *et al.*, 2011; Guglielmini *et al.*, 2011). For that reason, ICEs are considered as one of the main forces that drives the evolution of microbial genomes.

ICEs range from 18 Kbp (as in the case of Tn*916* of *Enterococcus faecalis*) to more than 500 Kbp (in the case of ICE*MlSym* of *Mesorhizobium loti*) (Johnson and Grossman, 2015) and generally code for adaptive genes that benefit the survival of their bacterial host. These cargo genes can mediate resistance to antimicrobials, bacteriophage infection or heavy metals (Burrus and Waldor, 2004; Balado *et al.*, 2013) and also can provide its host with N_2 fixation capability (Sullivan *et al.*, 2002), alternative carbon source utilization (Seth-Smith *et al.*, 2012) or the capacity to promote virulence or biofilm formation (Drenkard and Ausubel, 2002; He *et al.*, 2004; Davies *et al.*, 2009).

ICEs are structured in different functional modules: an excision/integration module; a DNA processing and transfer module similar to that of conjugative plasmids (see 1.1.1.3); a replication module; and a regulation module that controls the activity of the other modules. ICEs can also bear accessory modules encoding for the adaptive genes as described above, among others (Ghinet *et al.*, 2011; Guglielmini *et al.*, 2011).

The excision/integration process can be mediated by three different families of DNA integrases or recombinases: serine recombinases, DDE transposases or, most frequently, tyrosine recombinases(Wozniak and Waldor, 2010; Bellanger *et al.*, 2014). ICEs encoding tyrosine recombinases usually exhibit high specificity in *attB* recognition and integration. However, a few exceptions exist in which this protein can mediate integration at more than one attachment site (Song *et al.*, 2007; Doublet *et al.*, 2008; Sentchilo *et al.*, 2009; Wood and Gardner, 2015). The excision is often catalyzed with the help of a recombination directionality factor (RDF), also known as an excisionase, that changes the DNA architecture to displace the recombination reaction towards the excision of the ICE (Lewis and Hatfull, 2001).

Replication of excised ICEs prior to transfer to the recpient cell by RCR is a common step in the lifecycle of ICEs. It was first reported for ICEBs1 of *Bacillus subtilis* (Auchtung *et al.*, 2016) and since then, many other ICEs undergoing replication have been reported as, for example, Tn916 (Wright and Grossman, 2016) or ICESt3 of *Streptococcus thermophilus* (Carraro *et al.*, 2011, 2016), among others. Basically this replication uses the *oriT* as the origin replication and the relaxase along with other ICE and host-encoded elements as the replication initiator (Carraro and Burrus, 2014). Biologically, replication of the ICEs improves their stability as it prevents loss of the ICE in the case of chromosome replication while the element is excised (Lee *et al.*, 2010; Carraro *et al.*, 2015; Wright *et al.*, 2015).

In addition, accesory modules can code for maintenance systems, which prevent ICE loss in cell population. This is the case for SXT/R391 ICE of *Vibrio cholerae* whose srpMRC partitioning system segregates replicated ICE copies equally among the dividing cells (Carraro *et al.*, 2015). In addition, ICEs coding for Toxin-Antitoxin (TA) or Restriction-Modification (RM) modules could induce cell death or growth arrest upon ICE loss. While TA modules have been reported in ICEs as the mosAT system of SXT/R391 (Wozniak and Waldor, 2009), the role of RM systems in ICE loss prevention has not yet been reported even when they have been demonstrated as resistance factor against bacteriophage infection (Balado *et al.*, 2013).

Normally, ICEs remain as quiescent element integrated in the host chromosome. Active repression of transcription abolish the expression of the excision/integration and the transfer modules. Environmental stimuli can influence the expression and activity of the regulation module in the ICE and also of host-encoded factors which, in turn, can modulate ICE gene expression and transfer. The signals that induce ICE gene expression and the mechanisms of repression are not universally conserved, but there are some common observed features. For example, induction of SOS response can increase the transfer of SXT from *V. cholerae* by more than two orders of magnitude by repressing the repressor SetR (a homologue of the phage λ repressor CI) which inhibits the expression of an operon that includes the integrase and conjugation machinery activator genes *setD* and *setC* (Beaber and Waldor, 2004) (Fig. 1.7A). The same effect in transfer is produced in ICEBs1. In this case, the repressor ImmR inhibits the expression of the

excisionase and downstream genes. Upon induction of SOS response, the metallopeptidase ImmA cleaves the repressor (Bose *et al.*, 2008) leading to induction of the excisionase operon. Quorum sensing also plays an important part in the regulation of ICE transfer, as it is also the case of ICEBs1. The quorum-signalling peptide phospatase (PhrI) blocks the activity of response regulator aspartate phosphatase I (RapI) (Auchtung *et al.*, 2005). PhrI accumulates at high density of ICEBs1-bearing cells, leading to block RapI activity and consequently abolishing excision and transfer of ICEBs1. By contrast, in stationary phase, cultures containing a minority of ICEBs1-bearing cells lead to low levels of PhrI. Then, ICEBs1 transfer can occur. Finally, negative regulation produced by AbrB, a chromosome encoded transition state regulator, inhibits ICEBs1 excision. As a result of this regulation, transfer of ICEBs1 is abolished when the number of potential recipients is low (Auchtung *et al.*, 2005) (Fig. 1.7B).

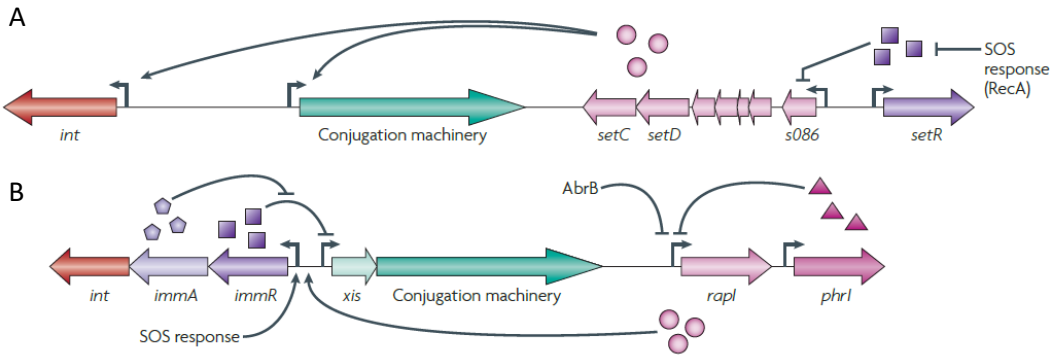


Figure 1.7. ICEs regulation modules: A: SXT ICE encodes the repressor *setR*, which represses *setC* and *setD* operon. B: ICEBs1 encoder the repressor *immR* and the antirepressor *immA*. Both repressors, upon induction of SOS response are cleaved, and proteins related to excision and DNA transfer are expressed. In ICEBs1 a second inhibition system exists, mediated by PhrI and AbrB which repress the expression of *rapI,* inhibiting excision and transfer of the ICE (Wozniak and Waldor, 2010).

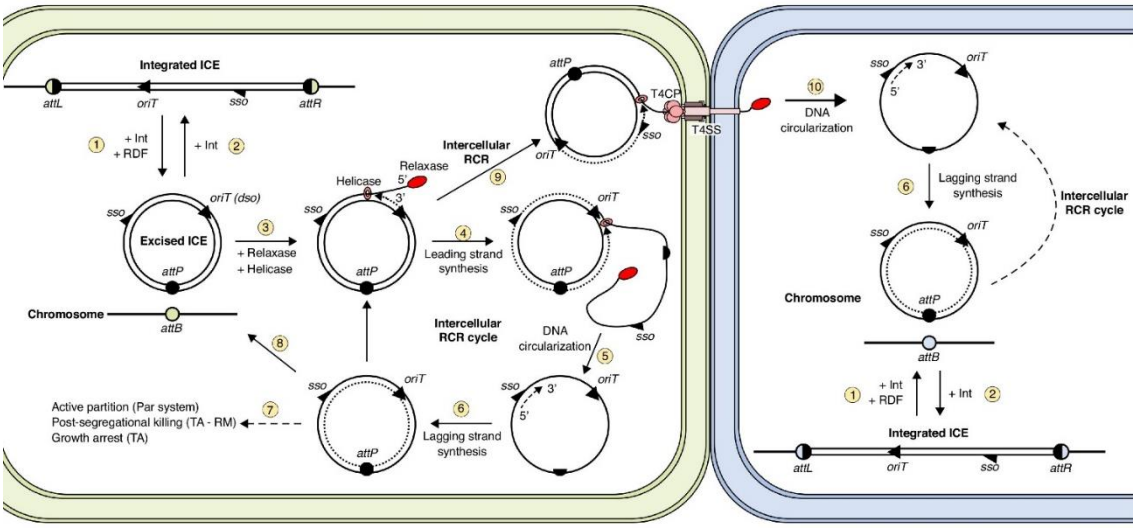


Figure 1.8. ICEs lifecycle: Excision (1) and integration (2) is mediated by an integrase helped by the excisionase or recombination directionality factor (RDF). Excised ICE serves as substrate for replication by rolling circle replication (RCR) (3-8) and for transfer to the recipient cell (3 and 9). This transfer is performed with ICE-encoded proteins. Once in the recipient cell, the DNA is recircularized (10) and lagging strand synthetized (6), followed by possible intercellular replication or integration in the chromosome (2). After replication ICEs can be stabilized (7), be reintegrated in the chromosome (8) or undergo multiple rounds of replication (Burrus, 2017).

Other MGE have been described to be integrated in the chromosome and even more widespread. This the case of Integrative and Mobilizable Elements (IMEs) that encode their excision/integration module but need to hijack the conjugation machinery of a conjugative element for their own transfer. These IMEs are detected alone in the chromosome, in tandem or inside ICEs. Another case is that of *cis*-Mobilizable Elements (CIMEs) elements that have lost their integration and transfer genes but retained *attL* and *attR* sites (Pavlovic *et al.*, 2004; Brochet *et al.*, 2008; Bellanger *et al.*, 2011, 2014). IMEs mechanisms to be maintained as an extra-chromosomal form have not been described yet. However, the study of the gene content of various IMEs suggests that the may use strategies (replication, addiction systems, partition) similar to ICEs (Guédon *et al.*, 2017).

1.1.3 HGT barriers

There are several prokaryotic defense strategies including, among others, RM systems, in which the restriction endonuclease cleaves mono or bicatenary DNA at specific

sequences and the methylase (mainly methyltransferases) prevents the cut by modifying the DNA (Tock and Dryden, 2005); CRISPR/Cas systems in which, briefly, the acquisition of short sequences from foreign DNAs (spacers) serve as substrate for targeting those same foreign nucleic acids by degradation mediated by the Cas proteins (Mojica *et al.*, 2005; Jinek *et al.*, 2012); and finally the interference mediated by prokaryotic Argonaute (pAgo), identified, at least in 9% of bacteria, which drives DNA-DNA interference, as reported in *Thermus thermophilus* and in archaea (Swarts *et al.*, 2014) (see 1.2.2).

1.2 The *Thermus* genus

The thermophilic genus *Thermus*, that belongs to the ancient *Deinoccoccus-Thermus* clade (Weisburg *et al.*, 1989; Omelchenko *et al.*, 2005), has been isolated from diverse natural and man-made thermal environments, including hot springs, gold mines, hot water pipes and hydrothermal vents, among others.

Thermus spp. are Gram-negative, non-sporulating, non-flagelated, slender bacillar shaped bacteria that tend to form septated filaments in exponential cultures on rich medium, being separated by binary fission when reaching stationary phase. Most strains are orange/yellow-colored due to the presence of a large fraction of carotenoid in their membranes (da Costa *et al.*, 2015).

They grow aerobically with high growth rates on complex medium at optimum temperatures ranging from 62 to 75^{o}C. Under anaerobic conditions, some strains can use nitrogen oxides or even metals as electron acceptors. They don't need specific amino acids or vitamins but some oligolements are essentials (Fe^{2+}, Mn^{2+}, Co^{2+} or Cu^{2+}) for growth (Brock, 1978).

Between all its species, *T. thermophilus* has a special importance as a model organism for basic and applied research. The isolates of this species grow easily under laboratory conditions and many of them preset natural competence, exhibiting high efficiency acquiring eDNA thanks to its NCA (Cava *et al.*, 2009).

Most of the strains of *T. thermophilus* are halotolerant, growing on media with yeast

extract and NaCl up to 6% due to the accumulation of trehalose and/or mannosylglycerate (Silva *et al.*, 2003; Alarico *et al.*, 2007). Besides these solutes, *T. thermophilus* presents an unconventional polyamine intracellular content, which includes long-chain and ramified-chain compounds which are not present in mesophiles and moderate thermophiles, besides the common ones (putrescine, spermidine and spermine). These complex and specific polyamines are essential for the synthesis of proteins and RNA and DNA stabilization at high temperature (Terui *et al.*, 2005; Oshima, 2007).

T. thermophilus cell envelope is composed by four layers: a cell membrane (IM), a thin layer of peptidoglycan wall covalently bound to a secondary cell wall polymer (SCWP) and the outer membrane which lays on a proteic S-layer (Cava, Pedro, *et al.*, 2004).

Currently, the genomes of several *T. thermophilus* strains are available online: HB27, HB8, JL-18, SG0.5JP17-16, ATCC 33923 and NAR1. Genome structure is preserved in the different strains, composed by a highly similar and syntenic chromosomes (1.2-2 Mbp), being this property employed for parenthood analysis in Chapter 8, and one or more megaplasmids. In the case of strains used in this work, HB27 harbors a megaplasmid of 0.23 Mbp (pTT27), while HB8 strain has highly similar pTT27 megaplasmid of 0.25 Mbp, another megaplasmid (pVV8) of 81 Kbp and finally a small one of 9 Kbp (pTT8). Genomes of this genus present a high content in G+C (around 60-70%), showing low gene duplication. Finally, at least, strains HB8 and HB27 have were shown to be polyploid, containing four to five and even seven to eight copies of the chromosome and megaplasmid/s at both exponential and stationary phases (Ohtani *et al.*, 2010; Li, 2019).

1.2.1 HGT in *Thermus* spp.

Thermus spp., exhibit very efficient horizontal gene transfer mechanisms. These mechanisms have been studied extensively in *T. thermophilus* strain HB27 and described in the following sections (reviewed in Blesa *et al.*, 2018).

1.2.1.1 Transformation

T. thermophilus presents a very efficient constitutive and non-inducible NCA, operative during all growth phases (Hidaka *et al.*, 1994) which can uptake up to 40 Kbp/s per cell in a promiscuous way since it can internalize DNA from different bacteria, archaea and even from eukaryotes (Schwarzenlander and Averhoff, 2006). NCA provides a high frequency of transformation in the HB27 strain. However, this efficiency varies greatly across the genus (Koyama *et al.*, 1986). This might be related to barriers involved in DNA acquisition (such as RM systems) that inhibit DNA incorporation despite the high conservation of NCA genes among strains.

In *T. thermophilus* HB27, 16 genes have been identified as required for the NCA (Averhoff, 2009), which have been assigned to three different groups (Fig. 1.9). Many of these genes have a role in T4P biogenesis (Friedrich *et al.*, 2001, 2002, 2003; Averhoff, 2009) which are also involved in twitching motility (Salzer *et al.*, 2014).

Briefly, the first group of genes includes conserved competence proteins. A ComEA homologue is located in the inner membrane and binds dsDNA, thereby contributing to the transport of eDNA through the IM (Salzer *et al.*, 2014). ComEC, also located in the IM, mediates DNA transport from the periplasm to the cytoplasm (Schwarzenlander *et al.*, 2009). DprA is suggested to stabilize eDNA and might be important for strands exchange during recombination as suggested by Yadav *et al.*, 2014 in *Bacillus subtilis*.

The second group is integrated by pilin-like proteins (PilA1-4), being required for both DNA uptake and pilus assembly (Friedrich *et al.*, 2003; Schwarzenlander *et al.*, 2009), a leader peptidase (PilD), an AAA-ATPase (PilF) essential for DNA transport through OM and polymerization of T4P (Schwarzenlander *et al.*, 2009; Rose *et al.*, 2011; Salzer *et al.*, 2013, 2014) and the rest of the proteins essential for T4P assembly (Friedrich *et al.*, 2002): an IM protein (PilC) which likely interacts with periplasmic and cytoplasmic proteins (Karuppiah *et al.*, 2010, 2013); PilM which forms a complex with PilN, interact with PilO and generates a transmembrane platform for the assembly of pilins (Karuppiah and Derrick, 2011) , and finally a secretin-like protein (PilQ) which forms homopolymeric complexes essential for the passage of DNA through the OM and for pilus formation (Schwarzenlander *et al.*, 2009; Burkhardt *et al.*, 2011).

The third group of genes is composed by ComZ which could be involved in DNA uptake through the OM (Schwarzenlander *et al.*, 2009); PilN and PilO whose role was described above and finally, PilW which is essential for both NCA and T4P in assembly and stability of the PilQ complexes (Rumszauer *et al.*, 2006).

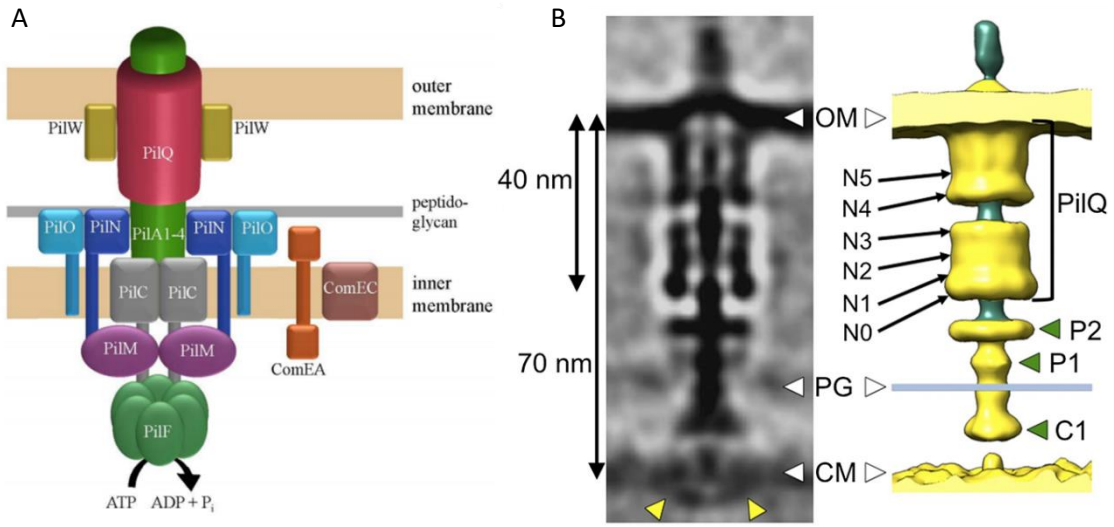


Figure 1.9. Natural competence apparatus of *T. thermophilus*. A. Structure: PilQ forms a channel through the OM and periplasm. PilW is essential for the assembly of the channel. Through this channel, a pilus-like structure formed by pilin-like PilA1-4 is polymerized with the help of PilF, the PilC dimer and also, by the interaction between PilO with a complex formed by PilM and PilN in the IM which serves as the assembly platform. Finally, ComEA binds incoming DNA and delivers it to ComEC in the IM (Blesa *et al.*, 2018). B. Subtomogram average on the left and 3D surface rendering on the right with the central pilus (green). N0-N5 indicates the putative domains of PilQ (Burkhardt *et al.*, 2012). Additional protein densities apart from PilQ are shown as green and yellow arrowheads. OM: Outer membrane; PG: peptidoglycan and CM: cytoplasmic membrane (Gold *et al.*, 2015).

1.2.1.2 Transjugation

T. thermophilus exhibits an unconventional mechanism for DNA transfer between cells that are in contact. This mechanism has been named "transjugation" (from transformation-dependent conjugation) and has been studied in the strain HB27. This mechanism produces higher transfer frequencies than other HGT mechanisms (Blesa *et al.*, 2017).

In this model, DNA is transferred in an active way from a donor to a recipient cell, in a process that can be bidirectional. In the donor cell, the role of the translocase TdtA (gene code *TTC1879*) is essential, as its mutation abolish transjugation completely. TdtA is a

membrane-associated protein of the FtsK family that assembles as a hexamer in the presence of ATP and shows ATPase activity. The 3D reconstruction of electron microscopy images suggest that TdtA can accommodate dsDNA, and sYFP (superfolder yellow fluorescent protein) fusions showed that it has a subpolar cellular location (Blesa *et al.*, 2017) (Fig 1.10A and B). Probably, accessory proteins collaborate with TdtA in this process. Unlike conventional models, the recipient cell has to participate actively in the transjugation process, as the functionality of the transformation machinery is essential for DNA uptake (Fig. 1.10C). In this process, any locus in the genome can be transferred to the recipient cell. However, loci in the megaplasmid show a 10-fold preference over loci in the chromosome. Furthermore, this process is not sensitive to DNA-DNA interference mediated by the Argonaute protein (TtAgo, *TTP0026*) in contrast to natural competence (see section 1.2.2) (Blesa *et al.*, 2015).

Some of the features of this mechanism are newly studied in this thesis.

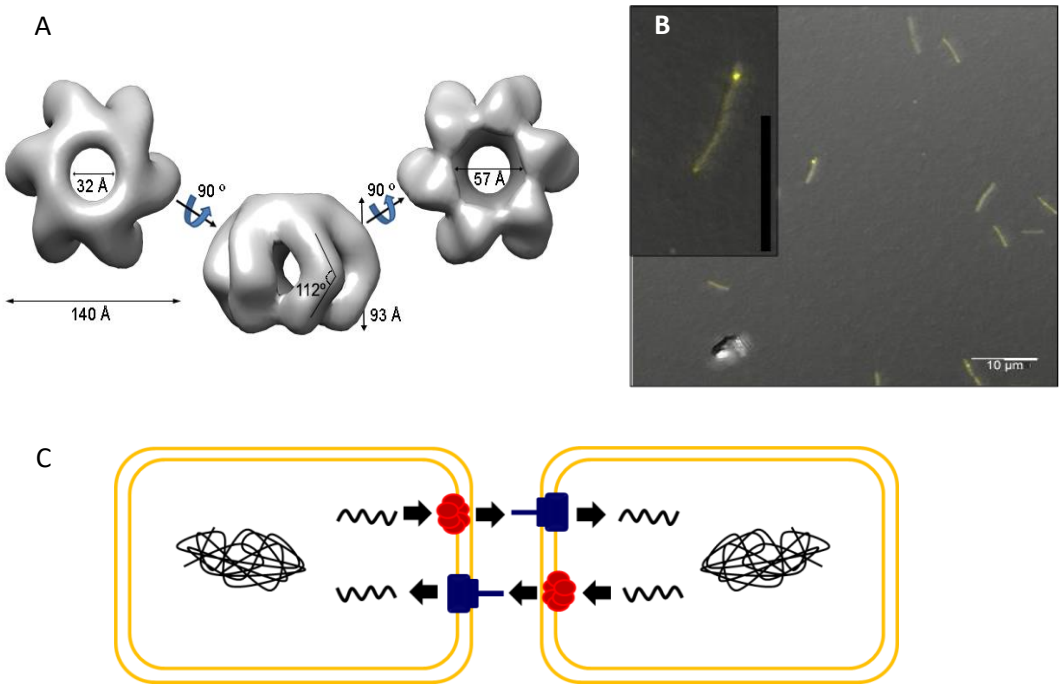


Figure 1.10: TdtA in transjugation. A. Three-dimensional reconstruction of TdtA by electron microscopy. B. Merging image of bright field and yellow channel of *T. thermophilus* expressing TdtA fused to sYFP from a multicopy vector (Blesa *et al.* 2017). C. Transjugation model: substrate DNA (represented as lines) is transferred from a donor cell through TdtA protein channel (in red) and uptaken into the recipient cell by the transformation machinery (in dark blue). Arrows represent DNA transfer direction and chromosome is depicted as a tangled line.

1.2.1.3 Transduction

This HGT mechanism has not been reported to date in *T. thermophilus.* In fact, most *Thermus* genomes so far sequenced do not present sequences corresponding to integrated prophages. This fact likely indicates that lysogenic phages in this thermophilic environment might not be common. Exceptions exist though. For example, *T. aquaticus* strain Y51MC23 contains two complete and two residual prophages (Brumm *et al.*, 2015) and *Thermus* strains RL, 2.9 and CCB_US3_UF1 contain a prophage each one (Dwivedi *et al.*, 2012; Teh *et al.*, 2012; Navas *et al.*, 2015).

However, the fact that infection caused by lytic prophages has been reported with several examples (Yu *et al.*, 2006) and the abundant presence of up to 3 CRISPR/Cas systems in this genus (Staals *et al.*, 2014) suggests the existence of transduction as a generator of variability in *Thermus* spp.

1.2.1.4 Vesicle-protected HGT

Transformation success depends on the length of the eDNA due to recombination requirements. Then, the length of eDNA in thermal habitats might compromise this mechanism, given the fact that high temperatures affect DNA stability (Soler *et al.*, 2008). In *T. thermophilus* and other *Thermus* spp., the production of vesicles containing DNA has been observed. These extracellular vesicles protect eDNA from nucleases allowing its displacement over long distances and seem to be a consequence of partial cell lysis during active growth, where DNA would be captured within membrane fragments (Blesa and Berenguer, 2015). EVs containing eDNA are produced in different *Thermus* isolates with sizes around 20 Kbp, being the whole genome represented. The production of these vesicles was higher at latency and exponential phase in *T. thermophilus* and practically undetectable when growing in minimum media (Blesa and Berenguer, 2015). How EV-associated eDNA enters the cell is not clear but depends on NCA, as mutants lacking competence genes are unable to acquire this DNA, but, as proposed for other organisms, vesicles might fuse to the OM or somehow eDNA becomes accessible for NCA (Blesa *et al.*, 2018).

1.2.2 Barriers against HGT in *Thermus*

In addition to the DNA restriction barriers that are widespread in bacteria, *Thermus* spp. present different innate defensive systems. On one hand, some strains of *Thermus* encode a homologue to the eukaryotic Argonaute protein (TtAgo) which reduces transformation efficiency at least one order of magnitude when eDNA is acquired by transformation (Swarts *et al.*, 2014; Blesa *et al.*, 2015). Entering eDNA is recognized by TtAgo loaded with complementary ssDNA guides and then, the target DNA is nicked. However, this mechanism, for an unknown reason does not affect DNA acquired by transjugation despite the involvement of the NCA (Blesa *et al.*, 2015). *In vivo* experiments corroborated that this is not due to G+C content (Blesa *et al.*, 2015) as suggested after *in vitro* assays (Swarts *et al.*, 2014, 2017).

On the other hand, 10-12 CRISPR arrays systems belonging to the IE, IIIa and IIIb families have been reported to date against invading DNA in *T. thermophilus*, being most of them located in the megaplasmid (Agari *et al.*, 2010; Staals *et al.*, 2014).

Finally, *Thermus* spp. encode different restriction-methylation systems which likely prevent from the entrance of invading DNA (Barker *et al.*, 1984; Wayne *et al.*, 1997; Zhu *et al.*, 2014).

Chapter 2

Scope and Objectives

Chapter 2: Scope and objectives

The work of this thesis focuses on the discovery and characterization of Integrative and Conjugative Elements (ICEs) described for the first time in *T. thermophilus*. The mainstay of this thesis is the knowledge of their lifecycle and the proteins involved in this process, with a special focus on the transjugation process. Therefore, the present thesis research work is oriented on the following objectives:

1. Describe bioinformatically the structure, organization and gene content of the ICEThs.
2. Analyze the lifecycle of the ICEThs, in terms of excision/integration, replication, transfer and maintenance.
3. Study the functionality of the different modules encoded in the ICEThs.
4. Explore the transjugation process, specially the rearrangement produced in the progeny as a consequence of this HGT process.

Chapter 3

Materials and Methods

Chapter 3: Materials and Methods

3.1 Materials

Strains used in this work are listed in Table 3.1. The plasmids used in this work are listed in Table 3.2, and Annex I describes the sequences of the oligonucleotides employed during this work. Finally, buffers and reagents used are grouped in Annex II.

Table 3.1. Bacterial strains used in this work.

Strain	Description/Genotype	Phenotype/use	Source
***Escherichia coli* strains**			
DH5α	*supE44 ΔlacU169 (Φ80 lacZΔM15) hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1 relA1*	Cloning	Hanahan, 1983
BL21 (DE3)	*F⁻ ompT gal dcm lon* HsdSB (r_B-m_B-) λ(DE3 [*lacI lacUV5-T7* gene 1 *ind1 sam7 nin5*])	Protein overexpression	Rosenberg *et al.*, 1987
***Thermus thermophilus* strains**			
HB27	*ATCC BAA-163/DSM7039*	Wild type (wt)	Y. Koyama
HB8	*ATCC 27634*	wt	Y. Koyama
HB27 Cm^r	HB27 spontaneous Cm resistant	Cm^R	This work
HB27 Δ*gdh*	HB27 Δ*TTC1211*::*kat*	Km^R. Chromosome labelled	Cava *et al.*, 2004
HB27 *ΔpilA4*	HB27^{EC} Δ*TTC0858*	Non-competent	Blesa *et al.*, 2015
HB27 *ΔpilA4* Δ*gdh*	HB27^{EC} Δ*TTC0858,* Δ*TTC1211*::*kat*	Non-competent, Km^R. Chromosome labelled	Blesa *et al.*, 2015
HB27 *ΔpilA4* 043 Δ*gdh*	HB27^{EC} Δ*TTC0858*, *hph17* insertion in 1,792,493 bp, Δ*TTC1211*::*kat*	Non-competent, Hyg^R, Km^R. ICETh1 labelled. Chromosome labelled	This work
HB27 043	HB27 *hph17* insertion in 1,792,493 bp	Hyg^R. ICETh1 labelled.	This work
HB27 *ΔpilA4* 061	HB27^{EC} Δ*TTC0858*, *hph17* insertion in 646,471 bp	Non-competent, Hyg^R, ICETh2 labelled.	This work
HB27 *ΔpilA4* Δ*attR1* 043	HB27^{EC} Δ*TTC0858*, 1.793.313 – 1,793,373::*kat* replacement, *hph17* insertion in 1,792,493 bp	Non-competent, Hyg^R, Km^R. ICETh1 excision blocked. ICETh1 labelled.	This work
HB27 *ΔpilA4* Δ*TTC0313* Δ*attR1*	HB27^{EC} Δ*TTC0858* Δ*TTC0313::hph5*, 1.793.313 – 1,793,373::*kat* replacement	Non-competent. Hyg^R, Km^R. Chromosome labelled.	This work

		ICETh1 excision blocked	
HB27 *ΔpilA4* Δ*int2* Δ*gdh*	HB27^{EC} Δ*TTC0858,* Δ*TTC0665::hph17,* Δ*TTC1211*::*kat*	Non-competent. Hyg^{R}, Km^{R}. ICETh1 and ICETh2 excision/integration blocked. Chromosome labelled	This work
HB27 *ΔpilA4* Δ*int2* 043	HB27^{EC} Δ*TTC0858,* Δ*TTC0665::kat, hph17* insertion in 1,792,493 bp,	Non-competent. Hyg^{R}, Km^{R}. ICETh1 and ICETh2 excision/integration blocked. ICETh1 labelled.	This work
HB27 Δ*exc2*	HB27 Δ*TTC0664::kat*	Km^{R}. ICETh1 and ICETh2 excision blocked.	This work
HB27 Δ*int1*	HB27 Δ*TTC1876::kat*	Km^{R}.	This work
HB27 Δ*int2*	HB27 Δ*TTC0665::kat*	Km^{R}. ICETh1 and ICETh2 excision/integration blocked	This work
HB8 Δ*int3*	HB8 Δ*TTHA1027*::*kat*	Km^{R}. ICEThs excision/integration blocked	This work
HB27 071	HB27 *hph17*::sIFP insertion in 1,792,493 bp	Hyg^{R}. sIFP reporter for ICEth1	This work
HB27 Δ*ttago*	HB27 Δ*TTP026*	Mutant in Argonaute	Swarts *et al.*, 2014
HB27 *ΔpilA4* *Δtth111II*::*kat*	HB27^{EC} Δ*TTC0858, ΔTTC1877*::*kat*	Non-competent. Km^{R}. Impaired in transjugation	Blesa *et al.,* 2017
HB27 *ΔpilA4* *ΔnurA::kat*	HB27^{EC} Δ*TTC0858, ΔTTC1878*::*kat*	Non-competent. Km^{R}. Impaired in transjugation	Blesa *et al.,* 2017
HB27 *ΔpilA4* *ΔtdtA*::*kat*	HB27^{EC} Δ*TTC0858, ΔTTC1879*::*kat*	Non-competent. Km^{R}. Deficient in transjugation	Blesa *et al.,* 2017
HB27 *ΔpilA4* Δ*methylase*	HB27^{EC} Δ*TTC0858, ΔTTC1880*::*hph17*	Non-competent. Hyg^{R}. Impaired in transjugation.	This work
HB27 *ΔppoL*	HB27 Δ*TTC0656*::*kat*	Km^{R}. Altered in ICEThs excision	Llamazares Master Degree

HB27 *ΔpilA4 ΔppoL* 043	HB27^{EC} Δ*TTC0858* Δ*TTC0656*::*kat hph17* insertion in 1,792,493 bp	Non-competent. Hyg^{R}, Km^{R}. Altered in ICEThs excision, ICETh1 labelled	This work
HB27 *ΔpilA4 ΔppoL* Δ*TTC0313*	HB27^{EC} Δ*TTC0858* Δ*TTC0656*::*kat* Δ*TTC0313*::*hph5*	Non-competent. Hyg^{R}, Km^{R}. Altered in ICEThs excision, Chromosome labelled	This work
HB27 Δ*toprim*	HB27 Δ*TTC0657*::*hph17*	Hyg^{R}. Altered in ICEThs excision	This work
HB27 Δ*TTC0658*	HB27 Δ*TTC0658*::*hph17*	Hyg^{R}. Altered in ICEThs excision	This work
HB27 Δ*peptidase*	HB27 Δ*TTC0663*::*hph17*	Hyg^{R}. Altered in ICEThs excision	This work
HB27 Δ*toxin*	HB27 Δ*TTC1785*::*hph17*	Hyg^{R}.	This work
HB27 *Δtth111II*	HB27 *ΔTTC1877*::*hph17*	Hyg^{R}. Impaired in transjugation.	This work
HB8 *ΔTTHA0672*	HB8 *ΔTTHA0672*::*hph5*	Hyg^{R}	This work
HB8 *ΔTTHB198*	HB8 *ΔTTHB198*::*hph5*	Hyg^{R}	This work

Table 3.2. Plasmids used in this work.

Plasmids	Description, use	Source
Minimal-ICETh1	pUC19:ICETh1attL:*int2:exc2:hph17*:ICETh1attR. *hyg* resistance cassette	This work
Minimal-ICETh1-YFRQ	Minimal-ICETh1 *int2:* Y364F, R331Q	This work
pET28b(+)	Km^{R,} *lacI*, T7 promoter. Protein overexpression vector for gene expression under T7 polymerase. Proteins can be fused to a 6xHis at N or C-terminal.	Novagen
pIB008	pUC19::*ΔTTC1880*::*kat.* Up and down arms for NurA mutation by insertion of *kat* resistance cassette	Blesa *et al.*, 2017
pIB009	pUC19::*ΔTTC1876*::*kat*. Up and down arms for Int1 mutation by insertion of *kat* resistance cassette.	This work
pIB039	pUC19::1,793,313 – 1,793,373 ::*kat* substitution. To eliminate ICETh1 *attR* by insertion of *kat* resistance cassette.	This work

pIB043	pUC19::1,792,493 *hph17* insertion. Up and down arms to insert *hyg* resistance cassette in ICETh1.	This work
pIB047	pUC19::*ΔTTC1877*::*hph17.* Up and down arms for *Tth111II* mutation by insertion of *hyg* resistance cassette	This work
pIB052	pUC19::*ΔTTC0665*::*kat*. Up and down arms for Int2 mutation. *kat* resistance cassette.	This work
pIB055	pMH 184 *hph17*::attBICETh1. pMH 184 hph17 harboring ICETh1 *attB1* site including 350 bp upstream and downstream. *hyg* resistance cassette. To detect ICEThs integration	This work
pIB059	pUC19::*ΔTTC1885*::*hph17.* Up and down arms for *TTC1885* mutation by insertion of *hyg* resistance cassette	This work
pIB060	pUC19::*ΔTTC0664*::*kat*. Up and down arms for excisionase2 mutation by insertion of *kat* resistance cassette.	This work
pIB061	pUC19::646,471 *hph17* insertion. Up and down arms to insert *hyg* resistance cassette in ICETh1.	This work
pIB062	pMH 184 *hph17*::attBICETh2. pMH 184 hph17 harboring ICETh2 *attB2* site including 400 bp upstream and downstream. *hyg* resistance cassette. To detect ICEThs integration	This work
pIB070	pUC19::*ΔTTC0657*::*hph17.* Up and down arms for *TTC0657* mutation by insertion of *hyg*resistance cassette	This work
pIB071	pUC19:: 1,792,493 sIFP and *hph17* both under individual PslpA promotor. *hyg* resistance cassette. *pslpA*:*sifp*:*pslpA*:*hph17* cassette extracted from pMH184 hph17 sIFP. To insert sIFP in ICETh1	This work
pIB077	pUC19::*ΔTTC0658*::*hph17.* Up and down arms for *TTC0658* mutation by insertion of *hyg* resistance cassette	This work
pIB078	pUC19::*ΔTTC0663*::*hph17.* Up and down arms for putative *TTC0663* by insertion of *hyg* resistance cassette	This work
pIB079	pUC19::*ΔTTHA1027*::*kat*. Up and down arms for Int3 mutation by insertion of *kat* resistance cassette	This work
pIB084	pUC19::*ΔTTC1880*::*hph17.* Up and down arms for *TTC1880* mutation by insertion of *hyg* resistance cassette.	This work

pIB085	pET28b(+) for *TTC1885* expression	This work
pIB087	pET28b(+) for *TTC1885* and previous ORF (putative antitoxin) expression.	This work
pLysS	Plasmid for expression of low levels of T7 lysozyme, an inhibitor of T7 RNA polymerase	Novagen
pMH184 hph17	Cloning vector for *T. thermophilus*. *hyg* resistance cassette.	Bosch *et al.*, 2019
pMH184 hph17(sIFP)	pMH184 hph17 intermediate cloning vector to obtain pIB071. sIFP extracted from pMotKpnqo(sIFP)	This work
pMotBlinkSEVA	Cloning modular vector for *T. thermophilus.* Bleomycin resistant.	Verdú *et al.,* 2019
pMotH	Cloning modular vector for *T. thermophilus*. *hyg* resistance cassette.	Verdú *et al.,* 2019
pMotH-SEVA	pMotBlinkSEVA. Replacement of Bleomycin for *hyg* resistance cassette between *AscI-PacI*. To measure transfer of a plasmid with *Tth111II* target.	This work
pMotH-SEVA-QX	pMotH-SEVA *hph17* thymine 435 changed by cytosine . To measure transfer of a plasmid without *Tth111II* target.	This work
pMotKpnqo(sIFP)	Modular plasmid pMotK encoding sIFP under Pnqo promoter. Km^R	Verdú *et al.,* 2019
pMot-Minimal-ICETh1	Minimal-ICETh1 *attL:exc2:int*2:*hph17*:*attR* cassette inserted in pMotH between PstI and EcoRI. To detect integration of Minimal-ICETh1 in HB8 strain	This work
pMot-Minimal-ICETh1-YFRQ	Minimal-ICETh1-YFRQ attL: *exc2:int*2*:hph17*:attR cassette inserted in pMotH between PstI and EcoRI. To detect integration of Minimal-ICETh1-YFRQ in HB8 strain	This work
pUC19	Cloning vector. Amp^R, P-lac-lacZ'	Vieira and Messing, 1982
pUC19::*TTC0313*::*hyg*	pUC19::*TTC0313*::*hph5*. Up and down arms for TTC0313 mutation in HB27 or *TTHA0672* in HB8 by insertion of *hyg* resistance cassette	Zafra *et al.*, 2002
pUC19::*TTP0146*::*hyg*	pUC19::*TTP0146*::*hph5*. Up and down arms for *TTP0146* mutation in HB27 or *TTHB198* in HB8 by insertion of *hyg* resistance cassette	Blesa *et al.*, 2015

3.2 Microbiological methods

3.2.1 Bacterial growth and storage conditions

Escherichia coli DH5α, used for cloning purposes, was grown at 37^{o}C in Luria Bertani (LB) medium (Lennox, 1955) under rotational shaking (180 rpm) in flasks filled up to 1/5 of total volume. Growth in solid medium was performed in LB medium containing 1.5% (w/v) agar at 37^{o}C. Plates supplemented with 5-bromo-4-chloro-3-indolyl-beta-d-galacto-pyranoside (X-gal 40 μg/ml) and isopropyl β-d-1-thiogalactopyranoside (IPTG 0.5 mM) were used for selection based in β-galactosidase activity.

Thermus thermophilus strains were grown at 60 or 65^{o}C in TB (*Thermus* broth) (see Annex II) under rotational shaking (180 rpm). For growth on solid medium agar 2 % (w/v) was added.

Selection of recombinant and transformed clones was performed with antibiotics. To that extent, media was supplemented with the appropriate antibiotic at the indicated concentrations: Kanamycin (Km, 30 μg/ml), Ampicillin (Amp, 100 μg/ml), Hygromycin B (Hyg, 100 μg/ml), Chloramphenicol (Cm, 20 μg/ml) or Bleomycin (3 μg/ml for *E. coli* and 15 μg/ml for *T. thermophilus*).

Monitoring of culture growth was performed with a Hitachi U-2000 spectrophotometer at OD_{600}.

Conservation of *E. coli* was performed in cryotubes supplemented with glycerol 20% (v/v) at -80^{o}C for long-term storage and at 4^{o}C on plates for short-term storage. In the case of *T. thermophilus*, short-term storage was at room temperature for plates and at -20^{o}C for cell pellets harvested at stationary phase for long-term storage.

3.2.2 Bacterial transformation

E. coli competence was induced through Inoue's method (Inoue *et al.*, 1990) and transformation was realized according to Hanahan's method (Hanahan, 1983). Highly competent cells of *E. coli* DH5α strain were grown in SOB medium and prepared according to $RbCl_2$ protocol with TFBI and TFBII solutions (Annex II).

Transformation of *T. thermophilus* was carried out by natural competence as described elsewhere (De Grado *et al.*, 1999). Briefly, mid-exponential cultures were transformed with 100 ng of DNA. After 4-hour incubation at 65^{o}C (60^{o}C in the case of *hph5*), cultures were spread on TB agar plates with the desired antibiotic and incubated for 48 hours in wet chambers to avoid desiccation. Transformants were restreaked at least twice on selective plates to homogenize the genotype of this polyploid bacterium.

Electro-competence of competence defective *T. thermophilus* strain was obtained as follows: an overnight culture was diluted 1/100 in TB and grown at 65^{o}C up to OD_{600}≈0.5. Cells were immediately transferred to ice and then centrifuged at 4^{o}C. The pellet was washed two times with 1/10 of the culture volume of glycerol 10% (v/v). Finally, cells were resuspended in glycerol 10% (v/v), aliquoted and stored at -80^{o}C.

E. coli BL21 strain, obtained from the fermentation facility of CBMSO (Centro de Biología Molecular Severo Ochoa), and *T. thermophilus* electro-competent (EC) strain were transformed by electroporation. 50 µl of competent cells with 10 ng and 0.1-2 µg of DNA depending on the use of *E. coli* or *T. thermophilus* strains, respectively. Then, the mix was placed into a cold electroporation cuvette (0.2 cm thickness, Bio-Rad Gene Pulser®). Cells were subjected to 5 ms electric pulse under a 12500 V/cm electric field (Equibio, Easyject Plus D2000; 2500V, 201 Ω and 25 µF). 600µl of pre-warmed media was added immediately and cells were incubated in 12 ml tubes at the appropriate temperature and time before plating.

3.2.3 Toxicity assays

To assay the toxicity of the putative ICETh1 encoded toxin (*TTC1885*) in *E. coli* BL21, three different strategies were used to limit its unwanted expression: pLysS plasmid was cotransformed along with the plasmid of interest; Antibiotic resistance induction and incubation were performed at 30^{o}C under low shaking (130 rpm); and glucose 0.2% (w/v) was added to medium (liquid and solid).

pIB085 and pIB087 were transformed in *E. coli* BL21 as reported above. Transformation frequencies were measured as the number of colony forming units (CFU) per µg of DNA used to transform the cells.

3.2.4 Transjugation assays

Transjugation assays were performed following the protocol described by Blesa *et al.,* 2015 with minor modifications. 200µl of saturated cultures of both parental cells were washed in one volume of TB and resuspended in 10µl of TB in presence of DNase I (5 units, Roche). The mixture was applied onto sterile 0.22µm nitrocellulose filters (GSWP, Millipore) and placed on top of prewarmed plates. Then, they were incubated at 60^{o}C for 4h 30min. Cells were detached from filters by vigorous shaking in TB and the desired dilutions were plated on selective TB plates at the same temperature. Transjugation frequencies were expressed as CFU of transjugants per CFU of recipient cells. Three independent assays were performed for each transjugation assay.

In the case of microscopy observations for sIFP marker transfer, overnight cultures were diluted to OD_{600}≈0.05 and grown on TB. At exponential growth phase, the recipient cells were covalently labelled at the surface with Texas Red succinimidyl Ester (Molecular Probes Europe BV, Leiden, The Netherlands) as described in 3.5.1. Then, mating assays were performed as described above. After incubation, cells were detached from the filter and processed as indicated in section 3.5.2.

3.2.5 Stress induction assays by UV light

Saturated cultures were diluted to OD_{600}≈0.05 and grown on TB. When cultures reached OD_{600}≈0.4, two mL of each culture were exposed to 60 J/m^2 of ultraviolet radiation (using an UV lamp Sylvania OSRAM StIII, Germany) for 20 min. Then both, UV exposed and control cultures were incubated at 65^{o}C with shaking for three hours in the absence of light. Cells were pelleted and total DNA extracted as indicated in 3.3.1.

3.3 Molecular methods

3.3.1 DNA manipulation

Total DNA from *T. thermophilus* cultures was extracted with DNeasy Blood & Tissue kit (Qiagen) following manufacturer's instructions for Gram-negative bacteria. RNase A (Roche) was added to the samples. A faster method was used for checking by PCR, based on the freeze-thaw technique (Miller *et al.*, 1999). Basically, 100 to 300µl of the culture

were centrifuged and resuspended in the same volume of Milli-Q® water. Samples were subjected to approximately 10 cycles of freeze by dried-ice followed by thaw at 42^{o}C for 3 min followed by vigorous vortexing. Finally, samples were boiled for 5 min at 98^{o}C, centrifuged and supernatant containing DNA was stored at 4^{o}C.

Plasmids were isolated using GeneJET Plasmid Miniprep Kit (Thermo Scientific) following manufacturers instructions.

DNA was digested with the appropriated restriction enzymes (FastDigest, Thermo Scientific) following manufacturers' indications. When needed DNA samples were, simultaneously, dephosphorylated using FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific) to prevent religation.

Digested DNA and PCR products were purified with Wizard® SV Gel and PCR Clean-Up System (Promega) according to manufacturer's protocol. DNA fragments were ligated using T4 DNA ligase in 10µl as indicated by the manufacturer (Promega). Ligation products were transformed in *E. coli* DH5α.

DNA concentration was measured with Nanodrop ™ One or ND-1000 (Thermo Scientific) spectrophotometer. PCR products were analyzed by electrophoresis at 100V in agarose gels (0.8-1.5% (w/v), low EEO, Conda Pronadisa) with 1X TAE (see buffers in Annex II) and stained with SYBR Safe DNA Gel Stain (Thermo Scientific). Plasmids were analyzed by restriction and both, plasmids and PCR products were sequenced when necessary at specialized companies.

3.3.2 RNA manipulation

RNA was isolated using RNeasy Mini Kit (QIAGEN) under manufacturer's instructions followed by DNAse I treatment (RQ1, Promega). RNA Integrity was checked either by agarose gel or with the Agilent 2100 Bioanalyzer.

Retrotranscription reactions (RT) were performed either with SuperScript III first strand synthesis kit (Invitrogen) followed by cDNA amplification using Pfu Ultra II Fusion HS DNA Polymerase (Agilent Technologies) or with the iScript cDNA Synthesis kit (Biorad PN170-8891), both under manufacturers' instructions and in both cases with random primers.

3.3.3 Mutants generation

Replacement mutants were obtained with plasmids indicated in Table 3.2. For their construction, pUC19 vector was used as backbone where two arms corresponding to 1 Kbp approximately of upstream and downstream regions of the target gene were cloned. Constructs were also designed as well to target intergenic regions without affecting gene expression (as in the case of ICEThs labelling). Between both arms, a gene cassette encoding a thermostable antibiotic resistance was cloned in the downstream direction (Figure 3.1). As the antibiotic resistance cassette does not harbor a transcription terminator, upregulation of downstream genes may occur. These constructs were transformed in the desired strain of *T. thermophilus* and positive clones, lacking the gene of interest and encoding the desired cassette were identified by PCR after two restreaking steps.

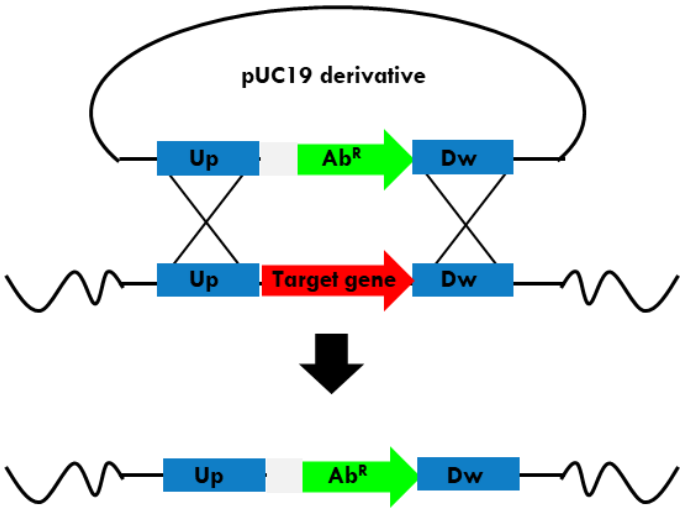


Figure 3.1. Replacement mutant strategy. A pUC19 derivative containing both up and down (dw) arms of a target gene recombines in the area with homology, producing a target gene-free organism resistant to the antibiotic resistance cassette (Ab^R) cloned in the plasmid. Note that pUC19 derivatives do not replicate in *T. thermophilus*.

3.4 Nucleic acids amplification

3.4.1 Conventional PCR

Conventional PCR was performed with different polymerases depending on the destiny of amplified DNA. For cloning, directed mutagenesis or cDNA amplification PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies) was used. However, for checking purposes GoTaq® Flexi DNA Polymerase (Promega) was employed. Oligonucleotides used for PCR purposes are indicated in Annex I.

Directed mutagenesis was used with pMot-Minimal-ICETh1 and pMotH-SEVA as templates to obtain derivatives pMot-Minimal-ICETh1-YFRQ and pMotH-SEVA-QX. Basically, conventional PCRs, but with 18 amplification cycles, with primers containing the desired mismatches (Annex I) were performed. Amplification products were incubated for 10 min at 37^{o}C with 1µl of *DpnI* FastDigest restriction enzyme. 5µl of the mixes were transformed in *E.* coli DH5α as reported in 3.2.2. Mutations were checked by sequencing.

3.4.2 qPCR (in collaboration with genomics and NGS facility at CBMSO)

DNA from three to five independent biological replicates was extracted as indicated in 3.3.1. Primers used are indicated in Annex I. In order to discard a potential contamination of reagents and/or primer-dimer artifacts a non-template control (NTC) reaction was carried out using all the reagents except the sample. For all tested genes NTC amplifications were always negative or delayed more than 5 cycles with respect to the experimental samples, which allowed to rule out contamination and primer-dimer artifacts. Standard curves were performed with a qPCR over an eight-point ¼ dilution curve made from the starting amplicon concentrations or from a "pool" sample: PCR efficiency is calculated from curve´s slope. Technical triplicates were performed in order to correct pipetting errors in plate loading.

Primers and mastermix: 1µl of primer mix (5 µM of each primer) + 5µl *Power* SYBR® Green PCR Master Mix (Thermo Fisher Scientific, CN 4367659) which includes AmpliTaq Gold® DNA Polymerase, dNTPs and the rest of reagents needed to perform

the PCR. As template of the problem samples 10ng of DNA was used per well. qPCR reactions were performed in a CFX384 Real Time System C1000 Thermal Cycler (Bio-Rad), in hard-Shell® 384-Well PCR Plates White Well Clear shell (Bio-Rad CN HSP-3805). Thermal conditions consisted of the following steps: 95 °C + (45´´x 95 °C + 45´´x 60 °C + 1´x 72ºC) x 40. Primer annealing/extension temperature and time was modified according to the amplicon size. We also included a *Melting* curve from 60 °C to 95 °C (0.5 °C/seg) at the end of the program to verify the specificity of the PCR. Fluorescence was acquired during both the 72 °C and *Melting* steps. The assay specificity for all tested genes was confirmed because all of them showed a unique *Melting* peak.

Relative quantification of the data was carried out using software GenEx v. 5.4.4 (MultiD

Analyses AB, Gothenburg, Sweden), performing the subsequent steps:

1- Efficiency correction of Cq (cycle quantification) values.

2- Average technical qPCR replicates.

3-For normalization the following quantification method was used: $2^{-\Delta(\Delta Cq)}$ (Livak and Schmittgen, 2001). 4-

Relative quantification ($2^{-\Delta Cq}$). ΔCq = Cq value of each individual sample against the calibrator sample (see below) Cq value.

First replica of each amplification product seen in every graphic was used as calibrator (=1) multiplied by the relative quantification number of each sample, to obtain a "normalized absolute quantification" data in every case.

Reference genes analyzed were 16S (*TTC3084*), DNA polymerase III (*TTC1806*) and RNA polymerase alpha subunit (*TTC1300*) only when normalization was applied. Suitable reference genes for normalization of data were evaluated by Cq stability and variability testing, using the GenNorm (Vandesompele *et al.*, 2002) and Normfinder (Andersen *et al.*, 2004) algorithms in GenEx v. 5.4.4. The geometric mean of the selected reference genes was used as reference index. The process of analyzing stability differs in both algorithms, so it may sometimes give different results. GenNorm algorithm was used to calculate the best pair of genes, usually with a similar profile. On the other hand, the NormFinder algorithm was used to order the genes by their individual stability value, as it denotes the number of genes with which the normalization would be more suitable.

Data absolute quantification processing was carried out using software Microsoft Excell 2010 (after correction of efficiency of Cq values in software Genex), performing the subsequent steps:

1- Standard curves data were plotted as a linear regression of the Cq values versus the log of the DNA quantities

2- Amplicon quantities from unknown samples were obtained by interpolating their Cq values into the standard curves

3- Quantification data were expressed as number of copies per ng of DNA.

In order to obtain the number of copies per ng of DNA, it should be noted that 10 ng of gDNA is added to every well.

3.4.3 RT-PCR (in collaboration with genomics and NGS facility at CBMSO)

RNA extraction and RT reactions were performed from three independent replicates as reported in 3.3.2 and the cDNA obtained was used as template. Same patterns and conditions were used as in qPCR but with a few changes. Master mix for *int1* consisted in: 1ul of primer mix (2.5 µM of each primer) + 5ul Sso Fast EvaGreen Supermix (Biorad, CN 172-5204) which includes Sso7d-fusion® DNA Polymerase, dNTPs and the rest of reagents needed to perform the PCR. PCR efficiency is calculated as well from a "pool" cDNA sample.

Thermal conditions were slightly different: 10´x 95 °C + (15´´x 95 °C + 1´x 60ºC) x 40 for all samples but *int1* that are 30´´x 95 °C + (5´´x 95 °C + 5´´x 60 °C) x 40.

gDNA contamination was assessed by PCR amplification of the samples without previous RT (Cqg). To the RT-PCR Cq values of each sample (Cq) were subtracted the corresponding Cqgs to obtain ΔCq for each sample. ΔCq were used to have a comparative estimation of the mRNA levels per cell of the different transcripts. The $2^{\Delta\Delta Cq}$ for each gene were obtained, subtracting the ΔCq values for each gene, to compare between two conditions, for example, a *recA* mutant with respect to the wild type or a UV-treated strain with respect to the untreated. Averages and standard deviations are given for all the values.

3.5 Processing of samples for microscopic visualization

3.5.1 In *vivo* fluorescent staining

1 ml of cells grown to exponential phase were centrifuged and re-suspended in the same volume of PBS. Then, labelled after 20 min incubation in the dark at 65^{o}C with 10µl of a 5 mg/mL solution of Texas Red succinimidyl ester in dimethyl sulfoxide. Addition of 100mM Tris-HCl pH 8 stopped the reaction. Finally, the cells were centrifuged and washed twice in PBS before being used in the transjugation experiments.

3.5.2 Sorting and confocal microscopy

Mated cells labelled with fluorescent (red) dye and the sIFP protein were sorted after washing with PBS and fixation with 1% (w/v) of paraformaldehyde.

Sorting was performed with FACSAria™ Fusion equipment (Becton Dickinson). sIFP was excited at 488 nm and emission was registered with 530/30 nm filter, while Texas Red was excited at 562 nm and its emission was recorded at 610/20 nm. Cell populations selected were laid onto microscope slides coated with 0.01% poly-L-Lysine, then covered with Mowiol-treated cover slips. Images acquisition was performed with Nikon A1R coupled to Eclipse Ti-E (Nikon) inverted microscope with 60x/1.4 oil Plan-Apochromat immersion objective under 488 and 561nm lasers to excite sIFP and Texas Red, respectively. Emission was registered at 525 and 595nm, respectively. Deconvolution of images was performed with Huygens 18.1.0p7 64b (S.V.I.) software and final design was carried out with Fiji software (Wayne Rasband, NIH, USA).

3.6 Transjugants analysis (in collaboration with Genomics and NGS facility at CBMSO)

After DNA extraction (3.3.1), WGS (Whole Genome Sequence) was performed by MicrobesNG using Illumina next-generation sequencing.

3.6.1 Assemblies and annotation sets

Three assemblies were provided by the sequencing center. Quality analyses were performed over reads using FastQC software (Wingett and Andrews, 2018).

3.6.2 Alignment and pre-process of alignment files

Reference genomes and plasmid sequences (*Thermus thermophilus* HB27 and HB8 strains), were downloaded from NCBI (Database resources of the National Center for Biotechnology Information., 2016). The reads of all samples sequenced were aligned against Thermus_thermophilus_HB27 and Thermus_thermophilus_HB8 reference genome using BWA aligner (Li and Durbin, 2009). We used Picard Tools to add that information and sort the alignment files. Several steps were followed according to Genome Analysis Toolkit (GATK) documentation (Garrison and Marth, 2012; Sandmann *et al.*, 2017; Weldatsadik *et al.*, 2017). The final obtained sorted BAM (Binary Alignment Map) file is then used in further analysis steps for Variant Calling.

3.6.3 Variant calling

The variant calling genotyping from pre-processed alignment files was performed using the GATK that offers a wide variety of tools with a primary focus on variant discovery and genotyping.

3.6.4 Identification of potential variants, filter, number and annotation of discovered variants

The first step was to identify potential variants on each sample using the HaplotypeCaller tool. By default, records were output even if they had very low probability of variation, in expectation that the VCF (Variant Call Format) will be filtered using tools such as vcffilter (https://github.com/vcflib/vcflib#vcffilter). The analyses were performed and results shown in Tables 8.2, 8.3 and 8.5. We performed the annotation of the discovered variants using snpEff software (Cingolani *et al.*, 2012). This software is a toolbox that performs genetic variant annotation, which annotates and predicts the effects of variants on genes. The annotation output files in the native VCF

format were processed using an in-house script written in Perl language (parser_gatkSnpeff_2xls.pl) to obtain a more readable, excel-like, annotation file.

3.6.5 Results analysis

In order to simplify those results, they have been processed using an in-house script written in Python language (compareSNP_betweenSamples.py) to compare all variants positions in samples to see positions that match between samples and to detect in which gene is located each variant.

3.6.6 Circos representation of variants against HB27 or HB8 genome and plasmid sequence

In order to select those transjugant variants against *T. thermophilus* HB27 or HB8 genome and plasmid sequences, we filtered the comparison variants position tables removing those variants present in HB27 reads against HB27 or HB8 reference sequences.

Then, Circos (Krzywinski *et al.*, 2009), a software package for visualizing data and information which let visualizes data in a circular layout, was used.

3.6.7 Resistance genes location: *kat* and *hyg* genes

The location of the resistance genes in the transjugants samples was identified with LAST software in order to align all genomes (references and sequenced genomes) against *kat* and *hyg* genes. To make the result more informative we aligned assembled genomes against reference *T. thermophilus* genomes and vice versa.

3.6.8 Strain-specific genes of donor strain in transjugant samples

In order to find strain-specific genes in the transjugants we generated a .fasta file with the sequences of all the genes of the HB27 and HB8 strains of *T. thermophilus*. Using BLAST (Basic Local Alignment Search Tool) command line of one against the other strain genome sequence, we obtained a list of genes specific for each strain. Then, using BLAST command line we checked for the presence of these genes within the sequence data of transjugants.

To avoid false positives, we filtered the BLAST alignment by 40 % of the query cover and used an e-value=0 to guarantee the identity.

3.6.9 Analysis of unmapped reads vs HB8 genome reference

Due to the low percentage of alignment reads from T1 to T8 sequenced samples against the HB8 reference sequence, all reads unmapped against the HB8 genome were extracted from all BAM files with seqtk tool (Robinson *et al.*, 2011), obtaining fastq files. Then, those files were parsed to fasta file with an in-house script in Python. In order to assign those unmapped sequences to a specific strain, we performed a Blast search analysis. Finally, Blast hits were transformed into their correspondent descriptions, and all frequencies related to the total hits of unmapped reads were calculated with an in-house script written in Python. The most of unmapped reads in transjugants T1-T8 corresponded to the chromosome and pTT27 megaplasmid of the HB27 strain, thus confirming the nature of the transjugants as HB27 derivatives that acquired a few DNA sequences from the HB8 strain (Chapter 8).

3.7 Bioinformatic toolbox

Sequence and ORFs analysis was performed with Snapgene® v. 4.2.11 (GSL Biotech LLC) and ApE plasmid editor (M. Wayne Davis). DNA sequence homology was executed with BLAST® (NIH). Multiple sequence alignments were executed using Clustal Omega (EMBL-EBI) (Madeira *et al.*, 2019). Protein homology was performed using with BLAST® (NIH) and UniProtKB database (Bateman, 2019). Protein domains were identified using the Pfam database (El-Gebali *et al.*, 2019). G+C content was calculated with the G+C Content Calculator (Biologics International Corp).

T. thermophilus genomes were browsed at BacMap server or at NCBI server. Existing ICEs were consulted using ICEberg database (MML, SJTU) (Liu *et al.*, 2018). tRNAs were assessed using tRNAscan-SE server (Lowe and Chan, 2016).

Graphics were represented using SigmaPlot v. 14.0 software indicating average and standard deviation. Statistical analyses were performed with SPSS ® Statistics v.25 (SPSS Inc., Chicago, IL, USA; 2008), considered statistically significant for p-values < 0.05 using student's t-test assuming homoscedasticity.

Chapter 4

Identification and bioinformatics analysis of ICEThs of *T. thermophilus*

Chapter 4: Identification and bioinformatics analysis of ICEThs of *T. thermophilus*

Three hypothetical Integrative and Conjugative Elements (ICETh1, ICETh2 and ICETh3) are identified and analyzed in this chapter. As *T. thermophilus* is only distantly related to other bacterial phyla, many proteins encoded in these elements have no homologues in the GenBank. For this reason, a table shows on each epigraph the most similar protein identified from any organism, excluding the clade *Deinoccoccus-Thermus*. Part of these results have been published in Blesa *et al.*, 2017 or have been accepted for publication (Baquedano *et al.*, 2019).

4.1 ICETh1 analysis

ICETh1 is a 14.857 Kbp long mobile genetic element flanked by 47 bp long direct repeats corresponding to the 3' end of an isoleucine tRNA (*TTC3049*) located at position 1,778,454 bp in the chromosome of *T. thermophilus* HB27. The G+C content of the element is 59%, 10 points lower than that of the chromosome. Domain and BLASTp (Protein Basic Local Alignment Search Tool) analysis of the element identified several ORFs that according to the putative function of their homologues were classified into different modules (Table 4.1). A putative tyrosine integrase (Int1, *TTC1876*) was classified as the only member of a hypothetical excision/integration module. A experimentally confirmed (Blesa *et al.*, 2017) DNA transfer module was identified, composed by a type IIG restrictase (*TTC1877*) identical to Tth111II (Zhu *et al.*, 2014), a NurA-like nuclease (*TTC1878*), the TdtA DNA translocase (*TTC1879*) and a hypothetical DNA methylase (*TTC1880*) preceded in the ICE sequence by the excision/integration module. Finally a third putative module (tentatively assigned as maintenance module) was identified that includes a putative type II Toxin-Antitoxin system, being *TTC1885* the HicA-like toxin and an upstream ORF without code number associated (corresponding to translation from 1,792,682 to 1,792,951 bp in the chromosome), the putative HicB-like antitoxin. Also, a putative phosphohydrolase (*TTC1884*), a complete DDE

transposase (*TTC1881*) and two more transposase fragments (*TTC1882* and *TTC1883*) have been also identified (Fig. 4.1).

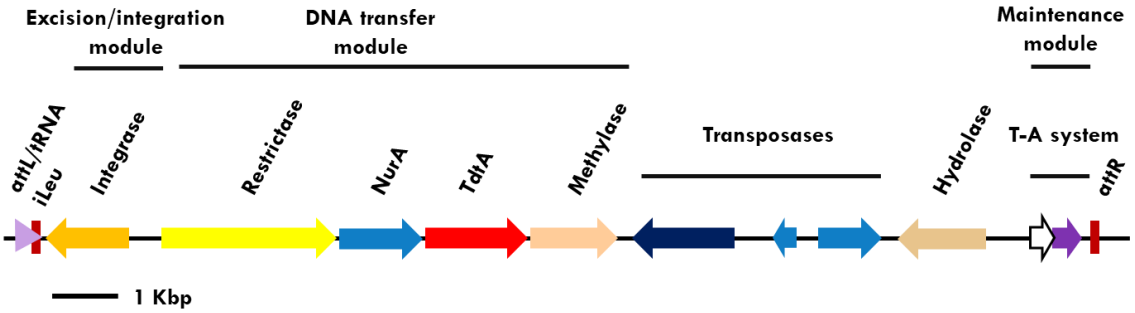


Figure 4.1. ICETh1 structure, showing genes encoded and putative functional modules. ORFs are drawn to scale and represented as arrows.

Gene	Description (ORF size)	Putative module	Domain	Homology (phylum)
TTC1876	Recombinase/ Integrase (392 aa)	Excision/ integration	Phage integrase	Site-specific recombinase (*Firmicutes*)
TTC1877	Restrictase Tth111II (1121 aa)	DNA transfer	N-6 DNA methylase	SAM-dependent DNA methyltransferase (*Chloroflexi*)
TTC1878	Nuclease NurA (377 aa)	DNA transfer	N/A	Single-stranded exonuclease (*Firmicutes*)
TTC1879	Translocase TdtA (568 aa)	DNA transfer	N/A	ATP-binding protein (*Acidobacteria*)
TTC1880	DNA methylase (410 aa)	DNA transfer	N/A	Modification methylase (*Firmicutes*)
TTC1881	DDE transposase (542 aa)	N/A	Transposase DDE	Transposase (*Cyanobacteria*)
TTC1882	Partial transposase (51 aa)	N/A	N/A	Transposase fragment (*Cyanobacteria*)
TTC1883	Partial DDE transposase (241 aa)	N/A	N/A	Transposase (*Euryarchaeota*)
TTC1884	Phosphohydrolase (427 aa)	N/A	HD domain	Hydrolase (*Aquificae)*
1792682-1792951 bp ORF	HicB antitoxin (putative) (89 aa)	Maintenance	N/A	Antitoxin HicB (*Spyrochaetes*)

TTC1885	HicA toxin (64 aa)	Maintenance	HicA toxin	YcfA family (*Cyanobacteria*)

Table 4.1. Description of genes encoded by ICETh1, putative modules in which they could be encoded, domain identification and closest match for protein sequence excluding *Deinoccoccus-Thermus* phylum. N/A means non-associated. aa refers to amino acids.

4.2 ICETh2 analysis

ICETh2 is even smaller than the previous element, with a total length of 11.276 Kbp located more than 1 Mbp away from ICETh1 in the position 641.867 Kbp. This new element is flanked by 47 bp direct repeats identical to the 3' end of a valine tRNA (*TTC3014*). The G+C content of this element (66 %) is more similar to the G+C content of the rest of the genome (69%). A similar approximation was performed with this element to assign ORFs to putative modules. We have identified a putative excision/integration module composed by a tyrosine integrase (Int2, *TTC0665*) and a excisionase (Exc2, *TTC0664*). Also we have identified, a putative replication module integrated by genes encoding the commercially available DNA primase/polymerase (PrimPol, *TTC0656*), used in the Trueprime^{c} (SYGNIS) amplification kit (Picher *et al.*, 2016), a homologue to eubacterial TOPRIM (topoisomerase-primase) domain of DnaG primases (Aravind *et al.*, 1998) (TOPRIM homologue, *TTC0657*) and other genes without homologues in the GenBank, which are likely cotranscribed along with the ones mentioned above (*TTC0654, TTC0655* and *TTC0658*). In addition, a putative protease (*TTC0663*) with a peptidase M66 domain was identified that could have a regulatory function (Fig. 4.2).

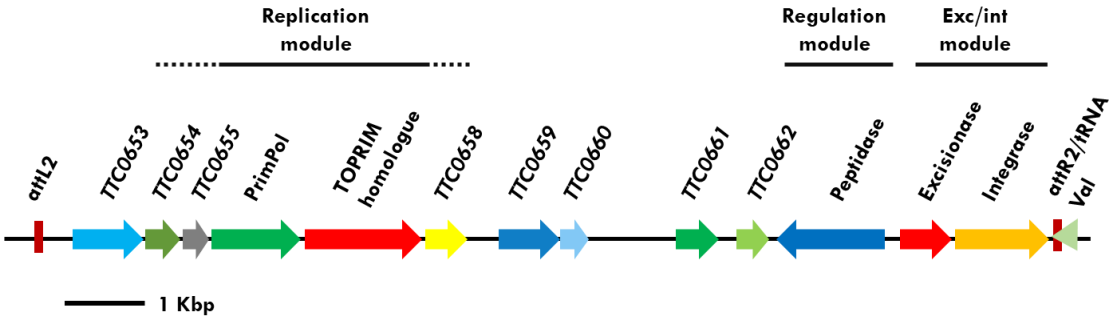


Figure 4.2. ICETh2 structure. The genes encoded by ICETh2 and its putative functional modules are depicted. ORFs are to scale and are represented as arrows. Dashed lines indicate uncharacterized ORFs that could be part of a putative functional module. Exc/int refers to excision/integration.

Gene	Description (ORF size)	Putative module	Domain	Homology (phylum)
TTC0653	Uncharacterized (221 aa)	N/A	N/A	No hit
TTC0654	Uncharacterized (94 aa)	Replication ?	N/A	No hit
TTC0655	Uncharacterized (64 aa)	Replication ?	N/A	MarR family (*Proteobacteria*)
TTC0656	PrimPol (293 aa)	Replication	Bifunctional DNA primase/polymerase	Bifunctional DNA primase/ polymerase (*Firmicutes*)
TTC0657	TOPRIM homologue domain protein (594 aa)	Replication	N/A	TOPRIM domain protein (*Proteobacteria*)
TTC0658	Uncharacterized (111 aa)	Replication ?	N/A	No hit
TTC0659	Uncharacterized (188 aa)	N/A	N/A	No hit
TTC0660	Uncharacterized (94 aa)	N/A	N/A	No hit
TTC0661	Uncharacterized (106 aa)	N/A	N/A	No hit
TTC0662	Uncharacterized (67 aa)	N/A	N/A	TonB dependent receptor (*Proteobacteria*)
TTC0663	Protease (522 aa)	N/A	Peptidase M66	Peptidase M66 (*Proteobacteria*)
TTC0664	Excisionase (155 aa)	Excision/ integration	Helix-turn-helix	DNA-binding protein (*Actinobacteria*)
TTC0665	Recombinase/ Integrase (391 aa)	Excision/ integration	Phage integrase	Site specific integrase (*Firmicutes*)

Table 4.2. Description of genes encoded by ICETh2, putative modules in which they could be encoded, domain identification and closest match for protein sequence excluding *Deinoccoccus-Thermus* phylum. N/A means non-associated. aa refers to amino acids. ? symbol refers to uncharacterized ORFs that could be part of a putative functional module.

4.3 ICETh3 analysis

In the related strain *T. thermophilus* HB8, ICETh1 and ICETh2 were not found. However, another putative ICE was detected at position 957,892 bp, integrated at the equivalent target were ICETh2 was integrated in the HB27 strain. Consequently, ICETh3 is flanked by 47 bp long direct repeated sequences corresponding to the 3' end of valine tRNA identical to the attachment sites of ICETh2. In this case, ICETh3 is larger than the ones in HB27, with a total length of 20,253 bp, and a G+C content of 67%, slightly lower than the chromosomal G+C content (69%). In this case, similar modules to the ones found for ICETh1 and ICETh2 in HB27 have been found, attending to the putative function of the proteins encoded within this new proposed mobile element (Table 4.3). In this case we have identified, i) an excision/integration module composed by a putative excisionase (Exc3, *TTHA1026*) and a tyrosine integrase (Int3, *TTHA1027*); ii) a putative replication module composed of genes similar to those in ICETh2, namely, a homologue to eubacterial primases (TOPRIM homologue2, *TTHA1022*) a helicase (*TTHA1020*) and probably *TTHA1023* as it seems to be cotranscribed with it and iii) a type II Toxin-Antitoxin system similar to ICETh1 composed by a HicA toxin (without code number associated) corresponding to the ORF encoded between position 958,105 and 958,299 bp in the chromosome and the putative HicB antitoxin (*TTHA1013*). The HicB antitoxin has been cristalyzed and its 3D structure resolved, but its function has not been described yet (Hattori *et al.*, 2005). Other hypothetical proteins have been detected such as an ATPase and another helicase (*TTHA1019* and *TTHA1017*, respectively). Finally, putative homologues to a transposase (*TTHA1018*), a nucleotidyltransferase (*TTHA1015*) and a methylase (*TTHA1016*) have been detected (Fig. 4.3).

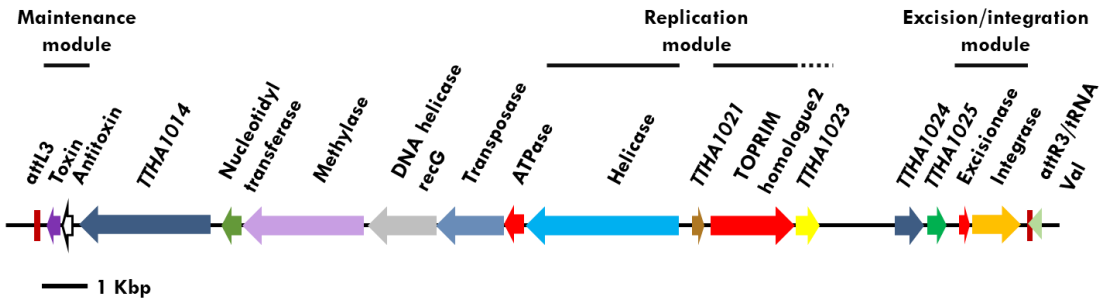


Figure 4.3. ICETh3 structure, showing genes encoded and putative functional modules. ORFs are to scale and are represented as arrows. Dashed line indicates uncharacterized ORFs that could be part of a putative functional module.

Gene	Description (ORF size)	Putative module	Domain	Homology (phylum)
958105-958299 bp orf	HicA toxin (64 aa)	Maintenance	HicA toxin	YcfA family (*Cyanobacteria*)
TTHA1013	HicB antitoxin (putative) (73 aa)	Maintenance	N/A	Antitoxin HicB (*Firmicutes*)
TTHA1014	Uncharacterized (893 aa)	N/A	N/A	ATPase (*Cyanobacteria*)
TTHA1015	Nucleotidyltransferase (98 aa)	N/A	Nucleotidyl transferase	Nucleotidyl transferase (*Proteobacteria*)
TTHA1016	Methylase (914 aa)	N/A	N/A	DNA methylase (*Cyanobacteria*)
TTHA1017	DNA helicase RecG (451 aa)	N/A	ATP dependent DNA helicase RecG Helix-turn-helix	ATP-depenedent DNA helicase RecG (*Proteobacteria*)
TTHA1018	Transposase (406 aa)	N/A	Transposase mutator family	Mutator family transposase (*Firmicutes*)
TTHA1019	ATPase (121 aa)	N/A	DNA-binding	AAA family ATPase (*Proteobacteria)*
TTHA1020	Helicase (1110 aa)	Replication	Helicase C-terminal SNF2 N-terminal	Helicase (*Cyanobacteria*)
TTHA1021	Uncharacterized (75 aa)	N/A	N/A	DNA-binding protein (*Firmicutes*)
TTHA1022	TOPRIM domain protein (592 aa)	Replication	N/A	TOPRIM domain protein (*Proteobacteria*)
TTHA1023	Uncharacterized (111 aa)	Replication ?	N/A	No hit
TTHA1024	Uncharacterized (260 aa)	N/A	N/A	No hit
TTHA1025	Uncharacterized (128 aa)	N/A	N/A	Leucin tRNA ligase (*Bacteroidetes*)
TTHA1026	Excisionase (59 aa)	Excision/integration	Helix-turn-helix	Excisionase (*Actinobacteria*)
TTHA1027	Recombinase/ Integrase (391 aa)	Excision/ integration	Phage integrase	Site-specific recombinase (*Firmicutes*)

Table 4.3. Description of genes encoded by ICETh3, putative modules to which they could be assigned, domain identification and closest match for protein sequence excluding *Deinoccoccus-Thermus* phylum. N/A means non-associated. aa refers to amino acids. ? symbol refers to uncharacterized ORFs that could be part of a putative functional module.

4.4 Comparison of putative ICEThs

The three putative ICEThs exhibit a common feature, the integrase (around 98% identity) (Fig. 6.1) and they also share homology in the terminal parts of the elements, around 1.2 Kbp in the integrase end of the element and around 140 bp at the other end. Interestingly, ICETh3 has common features with ICETh1 and ICETh2, like the Toxin-Antitoxin system (90% identity) and parts of the putative replication module (93% identity) (Fig. 4.4).

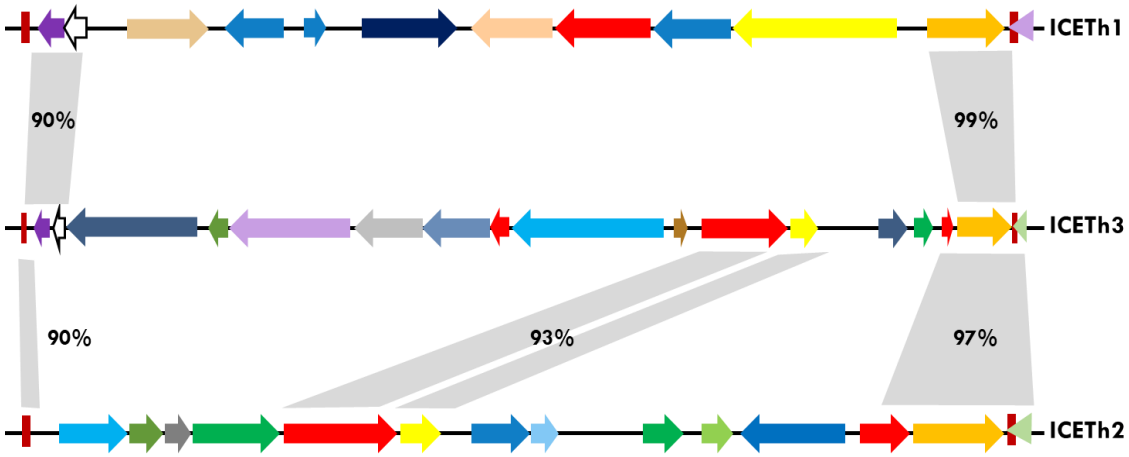


Figure 4.4. ICEThs sequence homology comparison. Related sequences areas are in gray with the nucleotide percentage identity indicated. Note that ICETh1 is represented in inverted orientation respect to figure 4.1.

The main ICETh1 and ICETh2 features and modules are analyzed through the following chapters. In the case of ICETh3, as it has been discovered recently, and in a different strain, only those features related with ICETh1 and ICETh2 are described in this thesis.

Chapter 5

ICEThs excision, transfer and copy number

Chapter 5: ICEThs excision, transfer and copy number

The bioinformatic analysis revealed the presence of three putative ICEThs, but some general characteristics had to be tested in order to determine if these elements actually behave as ICEs. In this chapter, excision of ICETh1 and ICETh2, transfer ability, and copy number have been studied.

5.1 Excision of ICETh1 and ICETh2 from the chromosome

To detect if ICETh1 and ICETh2 were able to excise from the chromosome, total DNA was extracted from 3 replicas of *T. thermophilus* HB27 cultures, at stationary phase, and PCRs were performed with convergent and divergent primer pairs whose combination allows to detect the integrated form of the ICE (*attL* and *attR*), the excised form (*attI*), and the "scar" left in the chromosome after excision (*attB*) (Fig. 5.1).

Figure 5.1. Scheme of ICE form detection. ICE is depicted as a dark red line and the adjacent chromosomal sites in black. Vertical lines indicates *att* sites. Blue arrows represent convergent primers and red arrows divergent primers to the mobile element.

In case of ICETh1, as seen in Figure 5.2, the integrated form was detected with primer pairs P306.3-P308 producing a PCR product of 566 bp for *attL1* and with primers P307-P309 for *attR1* that generate a PCR product of 1097 bp. Combination of these primers allowed us to detect the excised form *attI1* (primers P306.3-P307, generating a 1017 bp PCR product) and the *attB1* "scar" left in the chromosome, by recombination once the ICEth1 is excised (primers P308-P309, producing a 646 bp amplicon). The intensity of the *attI1* PCR product was lower than the rest of the products, meaning that this

amplification did not reach plateau phase. This fact suggested a lower copy number of the excised form of ICETh1.

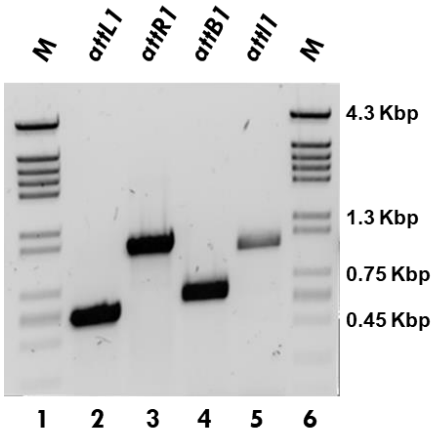


Figure 5.2. Agarose gel showing ICETh1 forms PCR products. Lanes 2 and 3 show the integrated form in the chromosome (*attL1* and *attR1* respectively). Lane 4 and 5 show the "scar" in the chromosome and excised form of the element (*attB1* and *attI1* respectively). PCR products were checked by sequencing.

To further characterize ICETh1, the relative proportions of the different forms were analyzed by quantitative PCR throughout the cell growth. To that extent, total DNA from three different cultures was extracted at five different times, from lag to stationary phase, and qPCR was performed. In this case, the primers used were different from those used in Figure 5.2. Primer pair P401-P402 detected the integrated form (*attL1*), oligo pair P338-P339 (*attB1*) was used to detect the "scar" in the chromosome and, finally, primers P401-P417 allowed for the detection of ICETh1 in the excised form (*attI1*).

Figure 5.3A reveals small differences in copy number of the different forms throughout the growth curve (one sample at lag phase, three through log phase, and one at stationary phase), being at stationary phase when the highest amounts of the excised form was detected. With this information, the copy number per ng of DNA extracted at stationary phase was analyzed with the same primer pairs. Results showed that the integrated form was around three orders of magnitude more frequent than the "scar" in the chromosome, while the excised form was ten times less frequent than the "scar" (Fig. 5.3B). Taken together, these data suggest that excision events are rare and once excised, ICETh1 does not replicate under our experimental conditions, in contrast to

conventional ICEs from other bacteria (Grohmann, 2010; Lee *et al.*, 2010; Carraro *et al.*, 2015).

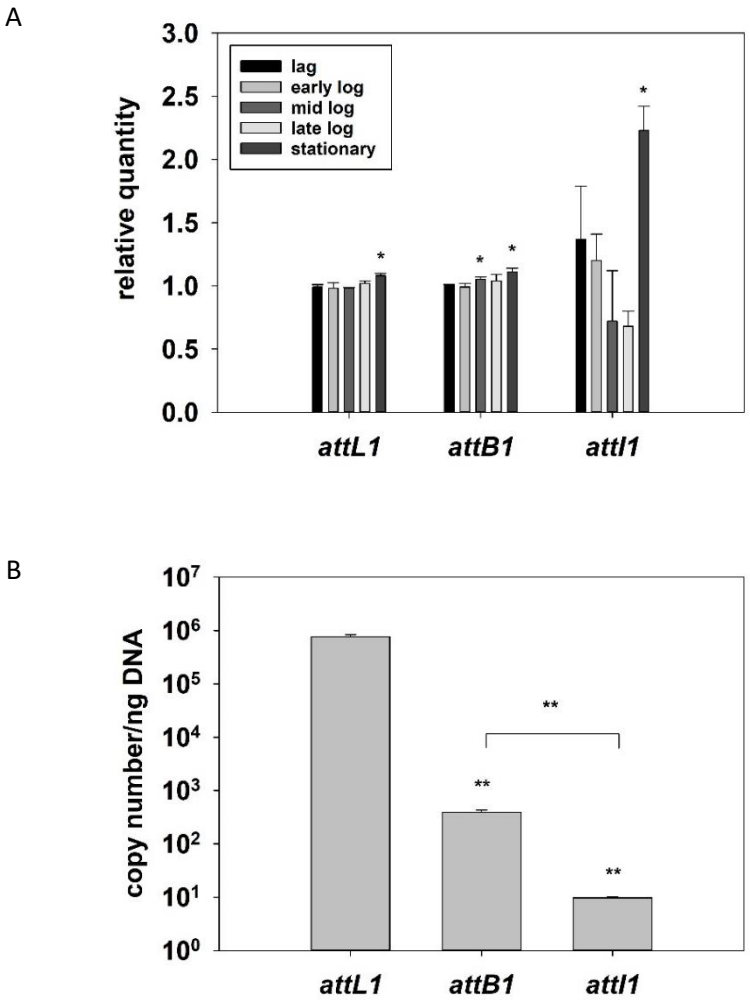


Figure 5.3. ICETh1 forms. A, relative proportion throughout cell growth and B, copy number per ng of DNA: integrated in the chromosome (*attL1*), chromosome resealed after ICE excision (*attB1*) and ICE excised (*attI1*). Samples for relative quantity are normalized with 16S (*TTC3084*) and DNA polymerase III (*TTC1806*) reference genes. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two). In case of A, comparison of each group with respect to lag phase. In A, the first replica of control cultures has been normalized to 1 on each form.

The same approach was used to study the excision of ICETh2 (fig. 5.4). Convergent and divergent primers to the element were designed as previously. This time, primer pairs P453-P454 and P455-P456 were used to detect *attL2* and *attR2*, producing PCR products of 362 and 290 bp, respectively. Excision (*attI2*) was detected with primer pair P488-P489, and the chromosomal "scar" (*attB2*) with primer pair P453-P456 generating PCR

products of 351 and 481 bp, respectively. In this case, for *attB2* and *attI2* plateau band detection, two consecutive PCRs were needed, suggesting, as in the case of ICETh1 a lower copy number.

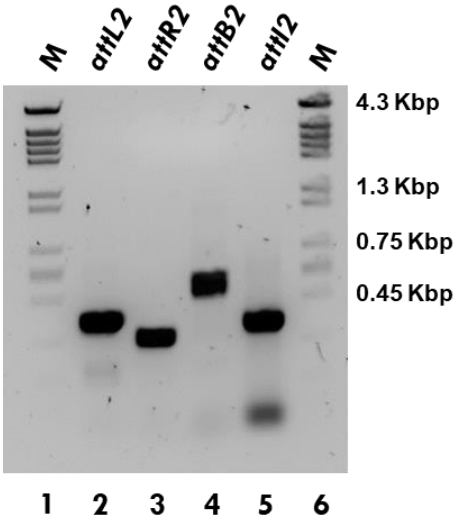


Figure 5.4. Agarose gel showing ICETh2 forms PCR products. Lanes 2 and 3 show the integrated form in the chromosome (*attL2* and *attR2* respectively). Lanes 4 and 5 show the "scar" in the chromosome and the excision of the element (*attB2* and *attI2* respectively). PCR products were sequenced.

Quantitative PCR of ICETh2 forms was performed as well with total DNA isolated from cells grown to stationary phase, using the same primer pairs of Figure 5.4, except that for *attB2* the pair P435-P438 was used (Fig. 5.5). Results are similar as in the case of ICETh1. The copy number of the integrated form of the element (*attL2*) is more than 4 orders of magnitude higher than the "scar" in the chromosome (*attB2)* or the excised form of ICETh2 (*attI2*). However, in this case, these two last forms, differing from ICETh1, are at very similar copy number. With this information, we were unable to determine if ICETh2 was active in replication.

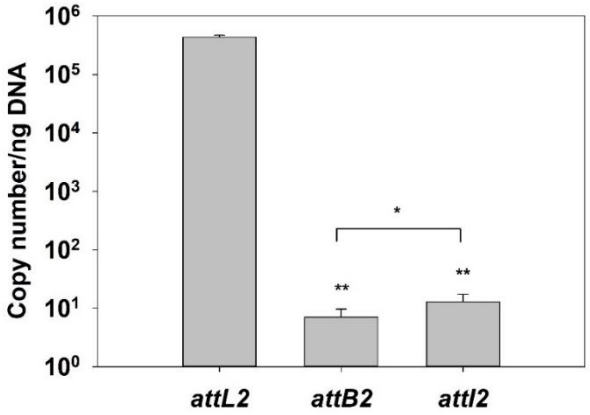


Figure 5.5. ICETh2 forms and copy number at stationary phase. *attL*2 refers to integrated form while *attB2* and *attI*2 refers to the chromosome "scar" and the circularized ICE, respectively. Samples are normalized with a PCR product amplified from the RNA polymerase alpha subunit (*TTC1300*). Asterisks indicate significant statistical differences between *attL2* and *attB2* or *attI2* (p-value<0.05, * and <0.001, **).

Finally, Argonaute protein (TtAgo) has been reported to be involved in cellular defense, more specifically in DNA-DNA interference. This interference takes place during transformation, decreasing its efficiency around one log, in which foreign DNA is cleaved by Argonaute. Plasmid copy number is decreased as well by this Argonaute mediated interference system (Swarts *et al.*, 2014). It has been previously reported by our group that DNA transfer by transjugation in *T. thermophilus* is insensitive to Argonaute-mediated interference (Blesa *et al.*, 2015).

ICETh1 and ICETh2 forms were measured in both wild type and Δ*ttago* strain in order to know if it was altered. DNA was extracted at stationary phase and quantitative PCR was performed with the same primers as above. As seen in Figure 5.6 ICETh1 shows a decrease in the "scar" (*attB1*) compared to its integrated form (*attL1*) slightly lower than in the wild type (wt). However, the difference between the "scar" (*attB1*) and the excised form (*attI1*) is slightly higher. In the case of ICETh2, it is noticeable that a difference between the "scar" (*attB2*) and the excised form (*attI2*) exists, reporting for the first time that excised ICETh2 shows a significantly higher copy number than the corresponding "scar" in the chromosome.

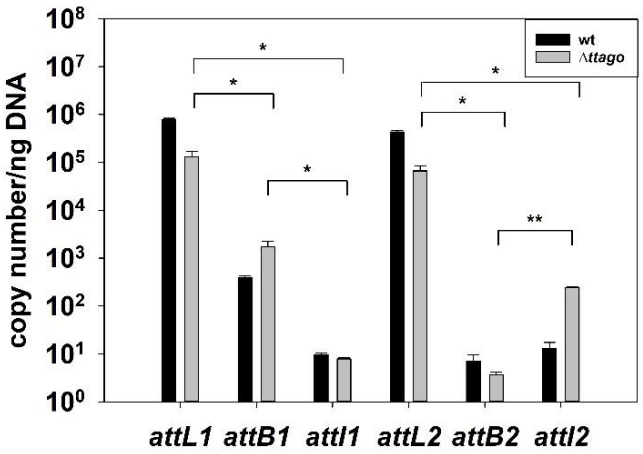


Figure 5.6. ICETh1 and ICETh2 form and copy number at stationary phase in wt (black bars) and Δ*ttago* strain (grey bars). *attL* refers to integrated form while *attB* and *attI* refers to excised form measured in the chromosome or in the ICE itself respectively. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two). (n=3). wt and mutant levels cannot be compared within the same ICE form, so comparison is made through difference of different forms within the same sample (wt or mutant).

5.2. ICETh1 and ICETh2 transfer to recipient cells

ICEs are usually transferred in the excised circular form, as conjugative plasmids. To characterize if this circular form is a better substrate for the transjugation machinery than other loci in the chromosome, the frequencies of transfer of both ICEThs was measured in comparison to a chromosomal marker. To that extent, two thermostable antibiotic resistance markers were used, Hygromycin B (*hyg*) in an intergenic region of ICETh1 (Δ*pilA4* 043) or ICETh2 (Δ*pilA4* 061), and Kanamycin (*kat*) in a different region of the chromosome corresponding to the glutamate dehydrogenase gene locus (*TTC1211*) (Δ*pilA4,*Δ*gdh*). All antibiotic resistance genes were introduced by double recombination after electroporation. To avoid bidirectionality produced during the transjugation process Δ*pilA4* mutants were used as donor cells, as they are unable to act as recipients (Blesa *et al.*, 2015). Spontaneous Chloramphenicol resistant (Cm^R) wild type cells were used as recipients. Parental cells grown to stationary phase were incubated together at 60^{o}C on nitrocellulose filters for 4h and 30 min. Selection of transjugants was perfomed according to the antibiotic resistance: Chloramphenicol/Hygromycin (Cm/Hyg) in the case of ICEThs transfer (Δ*pilA4* 043 or Δ*pilA4* 061) or Chloramphenicol/Kanamycin

(Cm/Km) in the case of chromosomal marker (Δ*pilA4,*Δ*gdh*). As seen in Figure 5.7, in spite of the apparent inability to replicate and its low excision frequency, ICETh1 transfer was significantly higher than the transfer of the chromosomal marker. ICETh2 was also transferred more frequently than the chromosomal marker but not as frequently as ICETh1. This suggests that a circular form is probably a better substrate for the transjugation machinery.

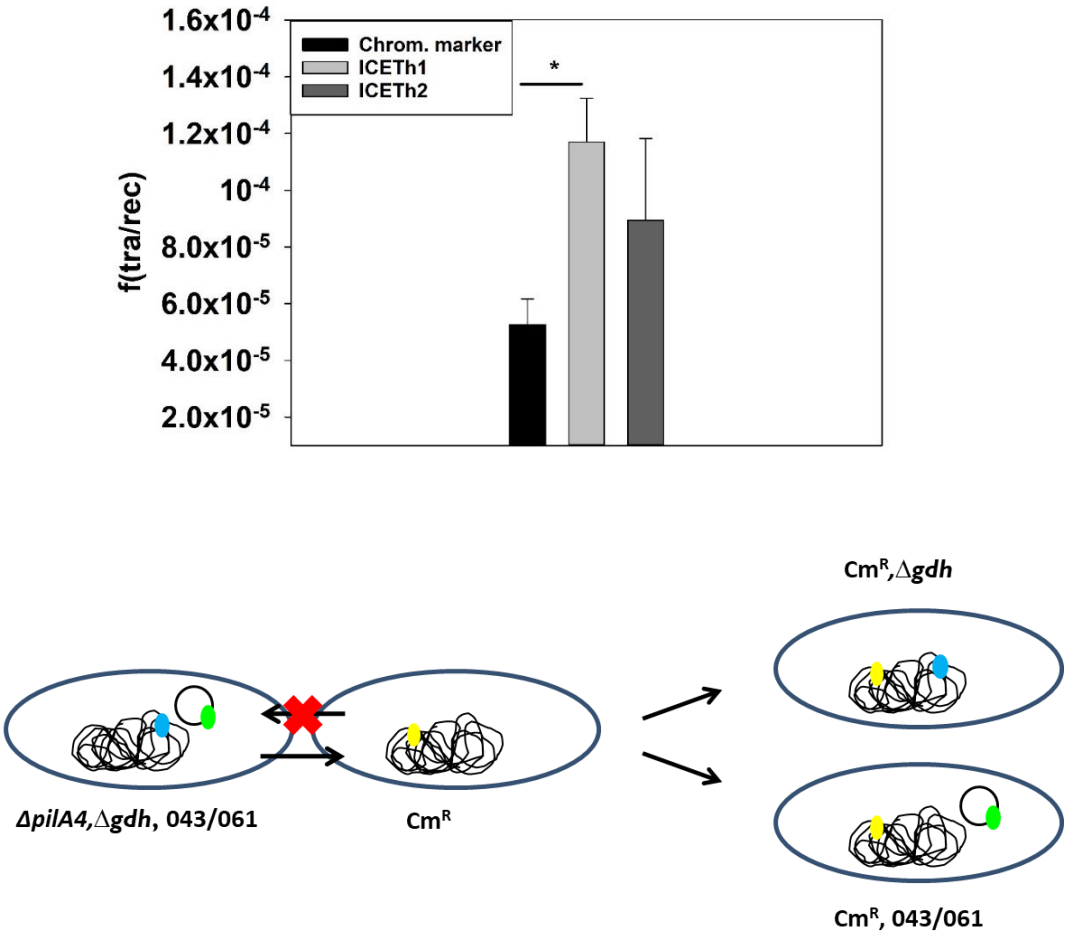


Figure 5.7. Transjugation frequencies of ICETh1, ICETh2 and a genomic locus analyzed as transjugants/ recipient cells. Transjugants were selected in Hyg and Cm plates in case of ICETh1 and ICETh2 while chromosomal marker was selected with Km and Cm. Asterisk indicate significant statistical differences (p-value<0.05). Below is depicted a scheme of the experiment: donor cell and recipient cell are depicted on the left and transjugants on the right. Both, ICETh1 and ICETh2 are shown at their excised form as a unique circle. Chromosome is drawn in black and different antibiotic resistances *(kat, hyg* and Cm^R) in blue, green and yellow, respectively. Red cross indicates that DNA cannot be transferred in the indicated direction.

In order to visualize the ICETh1 transfer, a thermostable version of GFP (Green Fluorescent Protein), named sIFP (Verdú *et al.*, 2019) was expressed in HB27 wild type

strain under the strong S-layer gene promoter (*PslpA*) integrated in the same location as the *hyg* as in the ICETh1 labelled 043 (HB27 071) (Fig 5.8A). These cells were incubated on a nitrocellulose filter along with *T. thermophilus* HB27 recipient cells previously labelled at amino residues of its surface with Texas Red succinimidyl esther (Acosta *et al.*, 2012). After incubation, cells were sorted in order to select those cells labelled with both green and red, and were visualized by confocal microscopy. Figures 5.8B and C shows unselected parental cells either in green (donor) or in red (recipient) and selected transjugants detected in yellow, thus showing the transfer of ICETh1.

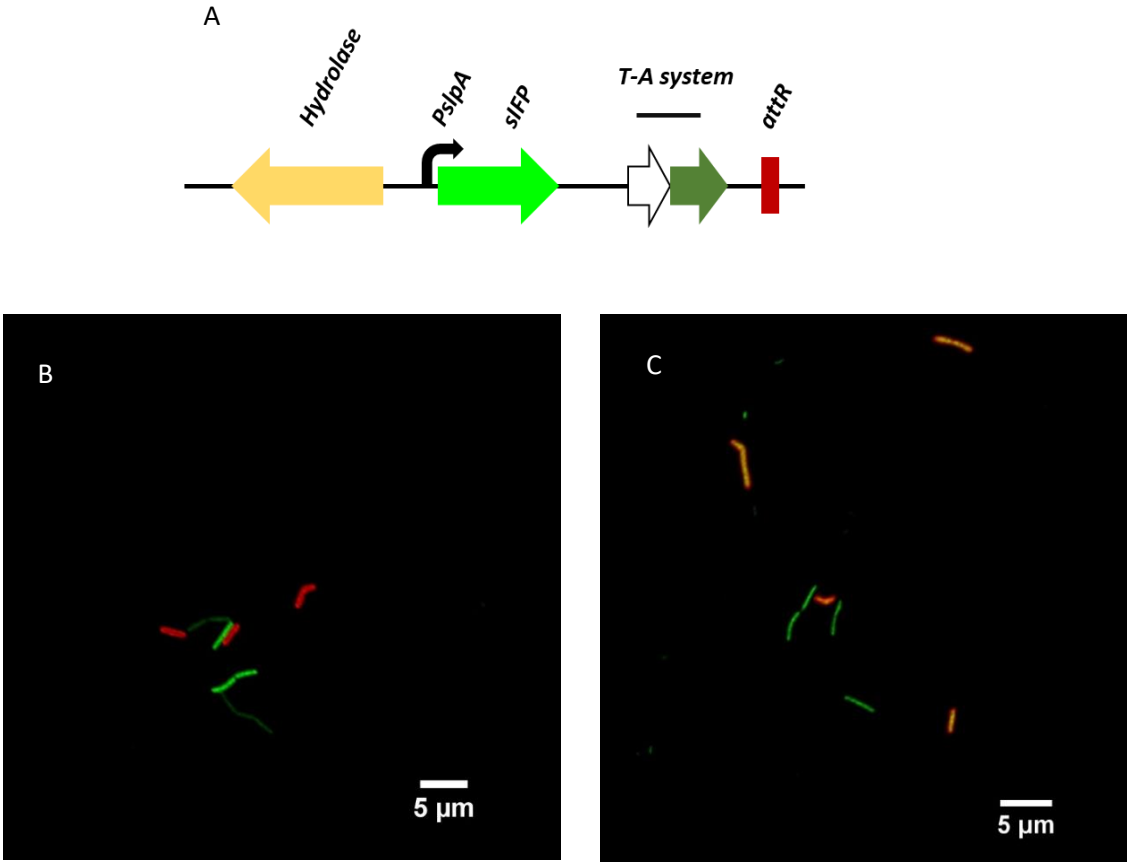


Figure 5.8. ICETh1 transfer. A. Scheme showing the sIFP integration site within ICETh1. B. Parental cells bound to Texas Red succinimidil ester in red, and cells expressing sIFP encoded within ICETh1 in green. C. Transjugants in yellow, labelled with both fluorophores and parental cells in green.

Chapter 6

Excision/integration and transfer modules

Chapter 6: Excision/integration and transfer modules

Once *T. thermophilus* ICEThs have been analyzed bioinformatically and their excision demonstrated in Chapters 4 and 5, the functionality of the putative modules for excision/integration and transfer, are tested in this chapter.

6.1 Excision/integration modules

As described in Chapter 4, two putative excision/integration modules have been found in *T. thermophilus* HB27. The first one composed only by Int1 (*TTC1876*), a putative tyrosine integrase, encoded within ICETh1 (Fig. 4.1). The second one, encoded within ICETh2, consists of Int2 (*TTC0665*), a putative tyrosine integrase 98% identical to Int1 (Figure 6.1), and Exc2 (*TTC0664*), a putative excisionase encoded upstream of Int2. Despite both Int1 and Int2 contains the catalytic residues of active integrases (Esposito and Scocca, 1997), the absence of an excisionase encoded in ICETh1, led us to check the functionality of both excision/integration modules.

Figure 6.1. Multiple sequence alignment between tyrosine integrases Int1, Int2 from *T. thermophilus* HB27 and Int3 from HB8 strain. Stars indicate conserved catalytic residues of tyrosine integrases. Red stars indicate residues mutated in Int2 for the isolation of pMot-Minimal-ICETh1-YFRQ plasmid variant (see below).

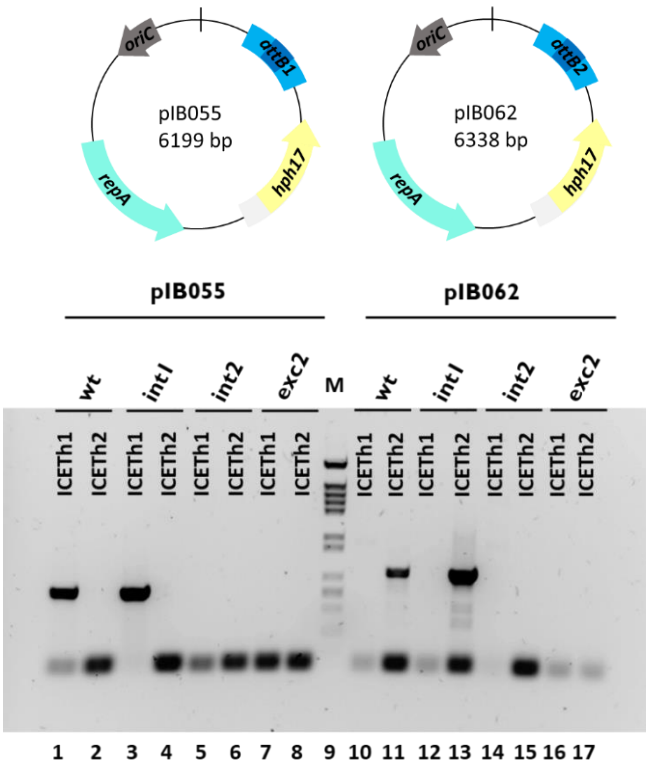
6.1.1. Mobilization of ICETh1 depends on ICETh2

To test mobilization, deletion mutants in the genes encoding Int1, Int2 and Exc2 were obtained (Δ*int1*, Δ*int2* and Δ*exc2,* respectively) by double recombination, in which the target genes were replaced by an antibiotic resistance cassette. In parallel, two replicative plasmids carrying *attB1* or *attB2* and their 350 bp flanking regions (upstream and downstream) were constructed as well (pIB055 and pIB062, respectively) as "*in trans*" insertion targets for ICETh1 and ICETh2, respectively (Fig. 6.2A). These plasmids were transformed in the wild type strain and in the three mutants. Due to the low frequency of ICEThs excision, transformants were selected after two rounds of growth on TB plates with the selection antibiotic. Then, total DNA was extracted from cultures grown to stationary phase, and the integration of both ICEThs in their respective

integration sites in the plasmids, was analyzed by PCR. For this, a common forward primer hybridizing in the plasmid (P451) and two reverse ones hybridizing specifically on each ICE (P482 for ICETh1 and P487 for ICETh2) were used. As seen in Figure 6.2B, each ICETh integrates at its specific *attB* site in the wild type strain (lanes 1 and 11), but no cross integration was detected (lanes 2 and 10). Also, the absence of Int1 didn't affect integration of either ICETh1 (PCR product of 592 bp, lane 3) or ICETh2 (PCR product of 781 bp, lane 13) in their respective targets . In contrast, the absence of Int2 or Exc2 abolished integration of both ICETh1 and ICETh2 in their respective *attB* sites (lanes 5, 7, 15 and 17). Thus, integration of ICETh1 depends on the Excisionase-Integrase module encoded by ICETh2. Noteworthy is the fact that previous versions of pIB055 and pIB062 harboring just 100 bp of flanking DNA of *attB* sites were not useful, as no integration of ICEThs could be detected. A similar result was obtained by quantitative PCR from the same samples (Fig. 6.2C). These data demonstrate that ICETh1 and ICETh2 mobilization (excision/integration) is mediated by the module encoded in ICETh2.

A

B



C

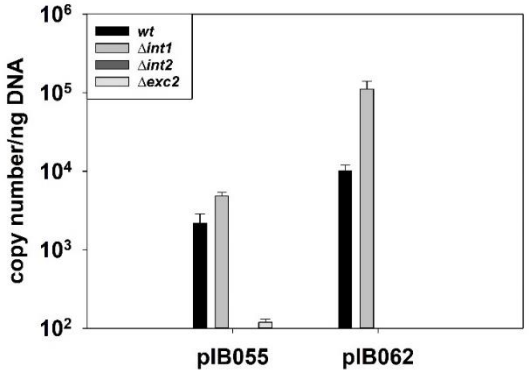


Figure 6.2. ICETh1 and ICETh2 integration detection. A. Replicative plasmids pIB055 and 062 harboring *attB1* and *attB2* and 350 bp of flanking DNA sites respectively. *hyg* resistance cassette (*hph17*) is in yellow under *slpA* promotor in light grey. In dark blue the 47 bp recognition site is shown while in lighter blue are the flanking areas to these sequences. *repA* encodes the protein necessary for plasmid replication in *T. thermophilus*. *oriC* is the origin of replication needed for plasmid replication in *Escherichia coli*. B. Agarose gel showing PCR products corresponding to ICETh1 or ICETh2 integration in plasmids pIB055 or pIB062 in different backgrounds: wt, *Δint1, Δint2* and *Δexc2*. Size markers (M) from top (bp): 4370, 2899, 2498, 2201, 1933, 1331, 1150, 759, 611, 453. C. quantitive PCR showing same features as B.

To confirm the functionality of ICETh2 excision/integration module on ICETh1 intracellular mobilization, a replicative plasmid containing the mobilization module was constructed. This new plasmid (pMot-Minimal-ICETh1) encodes the *exc2-Int2* module with their natural promoters, and a *hyg* resistance cassette for selection, flanked by the *attL1* and *attR1* sites (Fig. 6.3A). As a negative control of the integrase activity, a similar plasmid with mutations in two of the hypothetical catalytic residues of Int2 (Y364F and R331Q) was constructed (pMot-Minimal-ICETh1-YFRQ) (Esposito and Scocca, 1997). Both plasmids were transformed in the strain *T. thermophilus* HB8, an ICETh1 and ICETh2-free strain that contains the putative ICETh3 (see Chapter and Figure 4.3). As described in Chapter 4, ICETh3 encodes another putative tyrosine integrase (Int3) highly similar to Int1 and Int2 (98%) (Fig. 6.1) and an excisionase (Exc3) with 97% of identity to the one encoded in ICETh2 (Exc2). Hence, a deletion mutant for this protein (Δ*int3*) was also constructed to avoid any putative interference by this integrase. The pMot derivatives were also transformed in this mutant, and transformant cells striked twice on plates before analysis. Integration of the minimal ICETh1 encoded in the plasmids within its homologous *attB1* site (corresponding to iLeu-tRNA) was detected with primer pair P308-P514 producing a PCR product of 650 bp. As seen in Figure 6.3B, the minimal ICETh1 was able to integrate on its attachment site when, at least, any of the integrases

Int2 or Int3 were functional (lanes 1, 2 or 3). However, when none of the integrases was functional (Δ*int3* transformed with pMot-Minimal-ICETh1-YFRQ), there was no integration (lane 4). These data clearly demonstrate that Int2 from *T. thermophilus* HB27 and also the integrase Int3 from *T. thermophilus* HB8 can mobilize ICETh1, likely, in addition to their own ICE.

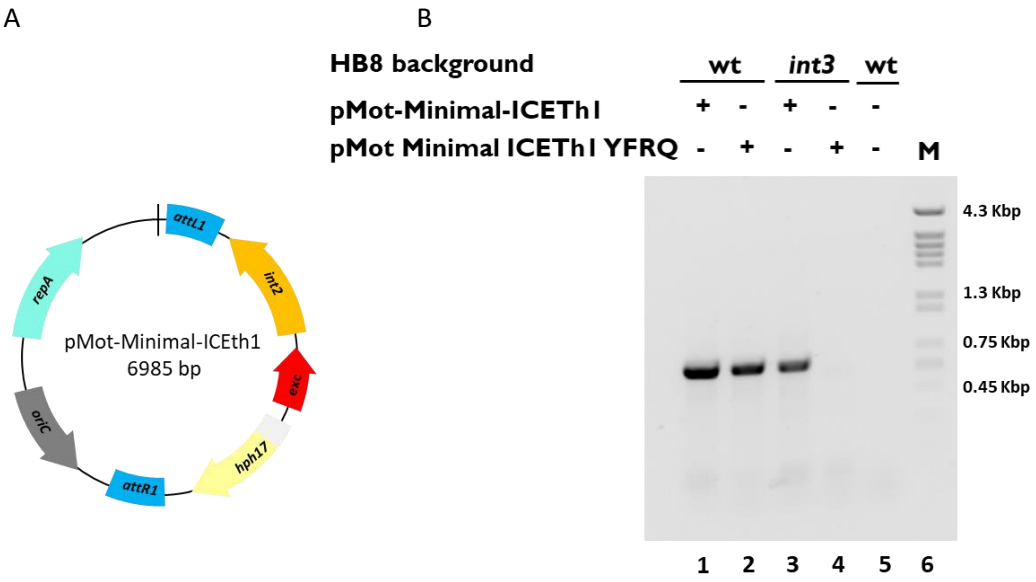


Figure 6.3. Minimal vector ICETh1 intracellular mobilization. A. Replicative plasmid pMot-Minimal-ICETh1 encoding *exc2* and *Int2* along with *hyg* resistance cassette between ICETh1 *attL1* and *attR1* sequences. B. Agarose gel showing PCR products that identify the integration of minimal ICETh1 in the genome of the wild type (wt) or *int3* mutant (*int3*) of this heterologous host. Presence of the indicated plasmid is depicted with + symbol and no presence with - symbol. YFRQ catalytically inactive mutant of Int2

6.1.2 *int1* and *int2* are transcribed at low levels

In order to know whether the futility of Int1 was due to a lack of expression, RNA was extracted from 3 wild type *T. thermophilus* HB27 cultures at exponential phase. This RNA was retrotranscribed and the expression of different genes was assayed with the appropriate controls. As seen in Figure 6.4 both *int1* and *int2* are expressed at low levels in comparison to other genes as *recA* (*TTC1644*) or the alpha subunit of ATP synthase (*TTC0907*) among others. The expression of *exc2* was assayed as well and exhibited similar levels to those exhibited by both integrases. These results indicate that the apparent inactivity of Int1 is not related to a lack of transcription. Other hypothesis are discussed in Chapter 9.

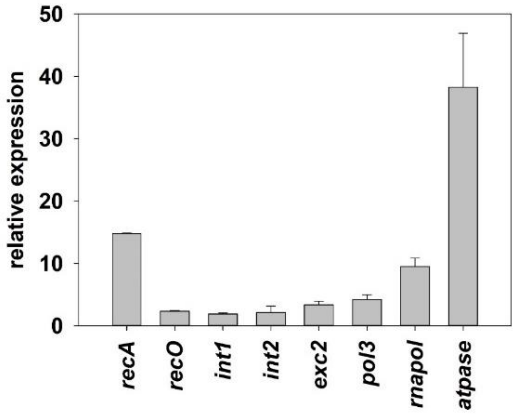


Figure 6.4. RT-PCR showing expression of different genes encoded in ICETh1 and ICETh2 compared with cellular housekeeping genes. pol3, DNA polymerase III (*TTC1806*); RNApol, alpha subunit of RNA polymerase (*TTC1300*); ATPase, alpha subunit of ATP synthase (*TTC0907*). Cq (cycle quantification) values from samples not treated with RT were subtracted from those treated with RT. Values correspond to averages from three biological samples.

6.1.3. Excision of ICETh1 affects to its intracellular mobilization

In order to know how abolition of ICETh1 excision from the chromosome affects transjugation, two mutants were constructed; in a Δ*pilA4* background, one lacking Int2 (Δ*pilA4, Δint2*), and the other a mutant in which the *attR1* site was replaced by *kat* resistance cassette (Δ*pilA4, ΔattR1*). Again, a spontaneous *T. thermophilus* HB27 Cm^R was used as recipient. As it is shown in Figure 6.5, transfer of the chromosomal marker was similar to that of the wild type in both mutants, whereas transfer of ICETh1 was clearly diminished, to a level that was below the chromosomal marker.

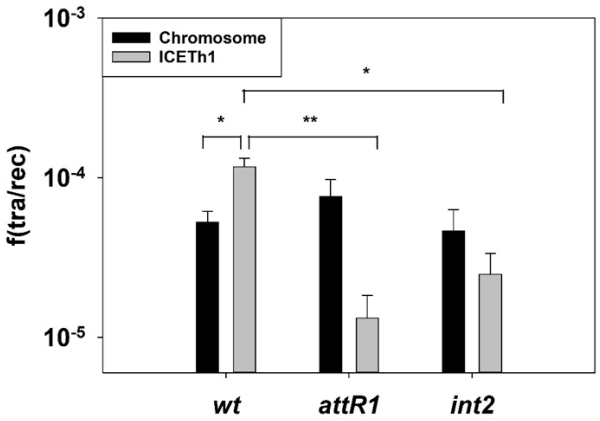


Figure 6.5. Transfer in *attR1* and *int2* mutants. Transjugation frequencies analyzed as CFU transjugants/CFU recipient cells of two antibiotic resistance markers (*kat* and *hyg*) located in the chromosome (black bars) or in ICETh1 (grey bars) in donor strains lacking the *attR1* site (*attR1*), the Int2 integrase (*int2*), or a wild type strain (wt). Cm^R was used to select recipient cells. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two).

6.2 Transfer module

The transfer module encodes the DNA translocase TdtA required for DNA donation in transjugation (A. Blesa, PhD thesis 2016). Here we show the results of the analysis of the role of the other genes encoded by this module, part of which have been published in Blesa *et al.*, 2017.

6.2.1 Transfer module is cotranscribed

As said in Chapter 4 this module would be composed by the restrictase Tth111II, NurA, TdtA and the putative methylase encoded downstream of TdtA. The distances between these genes suggest that they are cotranscribed. Only 4 bp separate *tth111II* and *nurA* and just 2 bp exist between this one and *tdtA*. Finally, *tdtA* sequence overlaps with the putative methylase encoded downstream. Genes upstream and downstream are encoded in opposite orientation. To confirm the expression of this operon, we used RNA from exponential cultures of *T. thermophilus* HB27 and checked by RT-PCR for the presence of intergenic transcripts with the primers indicated in Figure 6.6A. A conventional PCR was performed to confirm the absence of DNA in the RT negative reaction. As seen in Figure 6.6B, a common transcript was detected between *TTC1877-TTC1880* (lanes 2, 3 and 4) while no detection was obtained between *TTC1876-TTC1877* gene pair and *TTC1880-TTC1881* (lanes 1 and 5 respectively) confirming that the four genes are cotranscribed.

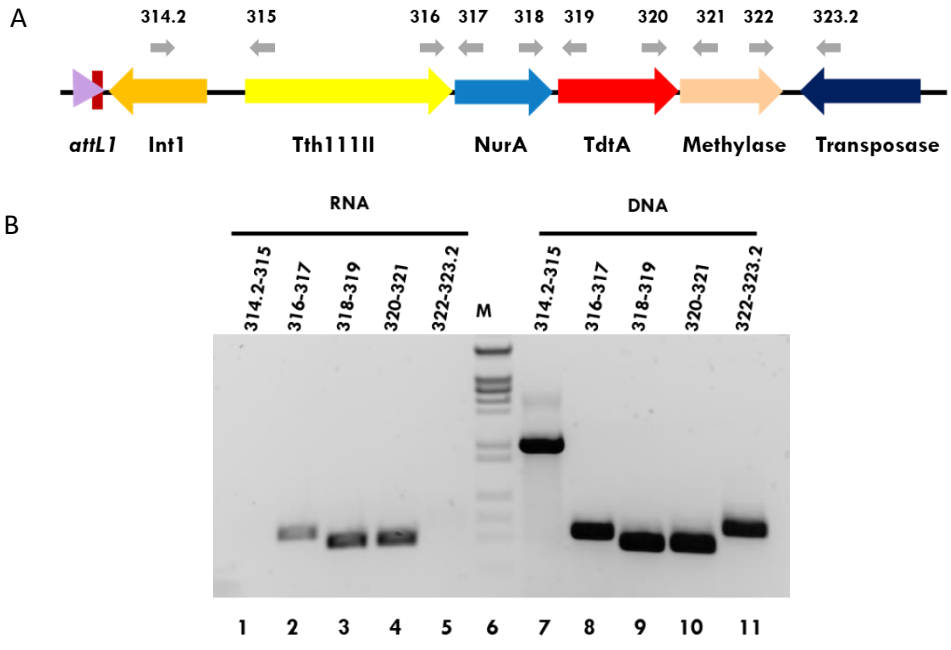


Figure 6.6. TdtA is encoded in an operon. A. Scheme of ICETh1 from Int1 to DDE Transposase. Grey arrows indicate the position where the primers for RT-PCR hybridize. B. Analysis by agarose gel of PCR products amplified with the indicated primers. Lanes 1-5 show reactions performed with RT-PCR as template. Lanes 7-11 show control reactions performed with DNA as template. Size markers (M) from top (bp): 4370, 2899, 2498, 2201, 1933, 1331, 1150, 759, 611, 453.

6.2.2 Mutants and their effects

In order to test the putative requirements for these genes in transjugation, we constructed knockout mutants in each of the genes in a Δ*pilA4* genetic background and tested their transjugation frequency. First, *Δtth111II* and *ΔnurA* mutants were tested. As controls, *Δgdh*, and *ΔtdtA*, both *ΔpilA4* were used. Deleted genes were substituted by *kat*. As recipient cells, we used a strain labelled with *hyg* in locus *TTC0313* (ferredoxin-nitrite reductase). The results shown in Figure 6.7 reveal that the absence of NurA produces a decrease in transjugation efficiency of 4-5 orders of magnitude with respect to the *Δgdh* control strain (bars 1 vs. 2). Similarly, absence of the restrictase Tth111II produces a decrease of around 3 orders of magnitude (bars 1 vs. 3) in the frequency of transfer. As expected in the *ΔtdtA* control the transjugation was completely abolished (bar 4).

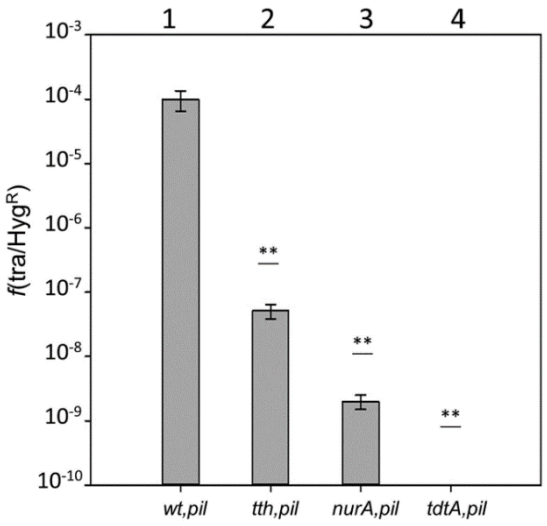


Figure 6.7. Transjugation frequencies of *Δtth111II* and *ΔnurA* analyzed as CFU transjugants/CFU recipient cells. *Δgdh* (wt) and *ΔtdtA* were used as controls. Transconjugants were selected in Km and Hyg plates. Asterisks indicate significant statistical differences compared to the wild type (p-value<0.001). (Blesa *et al.,* 2017).

The putative methylase (*TTC1880*) was mutated and analyzed in the same manner. In such a way that double mutants for the methylase and *pilA4* were used as donors. However, in this case, recipient cells were spontaneous wild type Cm^R cells and therefore, transfer frequency was measured for a marker located in ICETh1 (*ΔpilA4* 043) due to its higher frequency of transfer with respect to a chromosomal marker (as seen in Chapter 5). Figure 6.8 shows how transfer frequency decreased only twice with respect to the control, in contrast to the dramatic decrease produced by the other mutations.

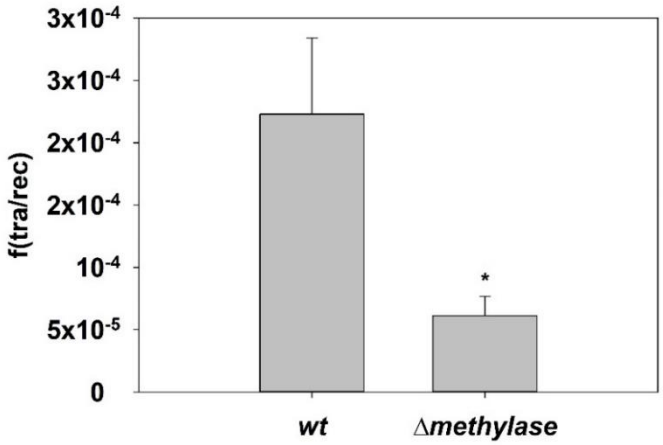


Figure 6.8. Transjugation frequencies of *ΔpilA4* Δ*methylase* mutant in comparison to *ΔpilA4,* marked within ICETh1 (wt). Analyzed as CFU transjugants/CFU recipient cells. Transjugants were selected in Hyg and Cm plates. Asterisk indicate significant statistical differences compared to the wild type (p-value<0.05).

6.2.3 Absence of a restriction site for transfer

Finally, in order to test if the presence of a *tth111II* recognition site is required for transfer, i.e. if it is recognized as *oriT* by the transjugation machinery, two similar plasmids were constructed (Fig. 6.9A). Both were modular replicative plasmids composed by the replication machinery for both *E. coli* and *T. thermophilus* and the *hyg* resistance cassette. *This hyg* has a target for the restrictase (CAARCA) so one of the plasmids has the wild type version of the gene (pMotH-SEVA) while the other introduces a single nucleotide change in the sequence (pMotH-SEVA-QX), preventing its recognition by the enzyme. This alternative version of the nucleotide sequence does not alter the amino acid sequence. Again, *ΔpilA4* background was used as donor cells (*ΔpilA4*,pMotH-

SEVA and *ΔpilA4*,pMoTH-SEVA-QX) and spontaneous Cm^R wild type cells were used as recipients. Transfer of both genes was measured and the result showed in Figure 6.9B. Transfer of the plasmid without *tth111II* target was very similar to transfer with the target site.

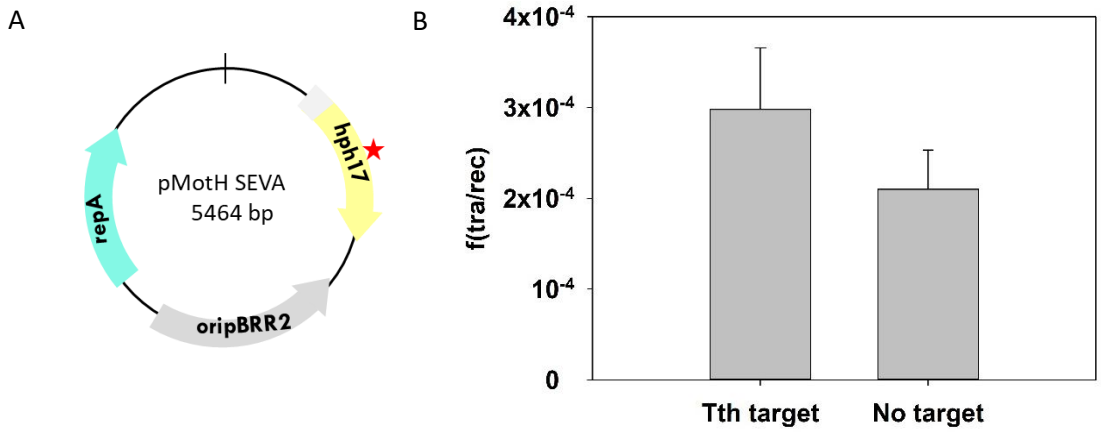


Figure 6.9. *Tth111II* target does not affect transjugation frequency A. pMotH-SEVA. *repA* encodes the protein necessary for plasmid replication in *T. thermophilus* and *oripBRR2* is used for replication in *E. coli*. *hph17* is the *hyg* resistance cassette. The red star indicates the position where *tth111II* restriction site is located. B. Transjugation frequencies of *ΔpilA4* pMotH-SEVA (Tth target) and *ΔpilA4* pMoTH-SEVA-QX (No target) analyzed as CFU transjugants/CFU recipient cells. Transjugants were selected in Hyg and Cm plates. No significant difference was found.

Chapter 7

Replication and accessory modules

Chapter 7: replication and accessory modules

Excision/integration and DNA transfer modules have been analyzed. In this chapter three putative modules, according to the bioinformatic analysis shown in Chapter 4 (replication, regulation and maintenance) are analyzed.

7.1 Replication module

Replication is a common feature in most ICEs (Grohmann, 2010; Lee *et al.*, 2010; Carraro *et al.*, 2015). However, in *T. thermophilus*, as seen in Chapter 5, replication seems unlikely for ICETh1, and, relatively poor, if at all, for ICETh2 under normal growth conditions. However, the presence in ICETh2 of genes that could be involved in replication suggests that we cannot discard ICEThs replication at least under particular circumstances. As mentioned above, this putative replication module would be formed by a DNA primase/polymerase (PrimPol) (*TTC0656*) and a protein with homology to domains of the eubacterial primases (TOPRIM homologue, *TTC0657*). In addition, other ORFs (such as *TTC0654, TTC0655* or *TTC0658*) which could be cotranscribed with them may also be involved in this process. Interestingly, *TTC0657* and *TTC0658* are conserved in ICETh3.

7.1.1 ICEThs excision is promoted by stress.

In order to know if stress conditions could promote replication, *T. thermophilus* HB27 wild type cultures were grown to exponential phase (OD_{600}≈0.4) and subjected to a 20 min treatment of UV light. Then, both treated and untreated control cultures were grown for 3 hours in the absence of light and total DNA was isolated and analyzed by qPCR to quantify *attL*, *attB* and *attI* forms of ICETh1 and ICETh2. Results are shown in Figure 7.1. In the case of ICETh1 (Fig. 7.1A), the excised form increased by 3-fold respect to control (normalized to 1 on its first replica) whereas no difference was detected for the "scar" *attB1*. On the other hand, a minor decrease in *attL1* was detected. In the case

of ICETh2 (Fig. 7.1B) the excised form (*attI2*) was 7-fold more abundant in cultures treated with UV in comparison to untreated control (normalized to 1). Contrary to ICETh1, *attB2* shows the same increase pattern as *attI2* with respect to untreated cultures. Therefore, the most likely explanation for these results is an increase in the excised forms of both elements under UV-induced stress condition, that could imply a higher excision rate or replication.

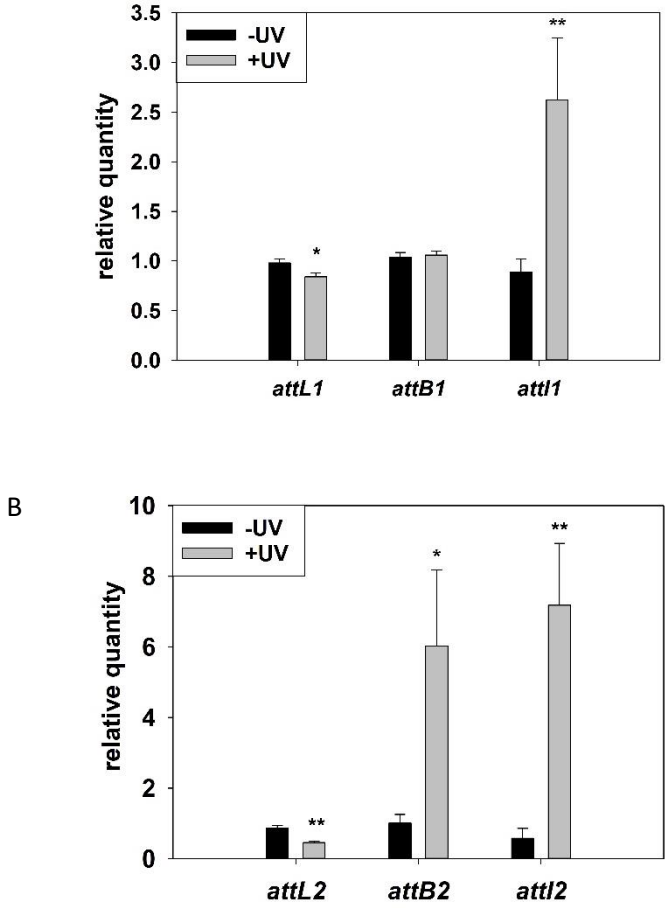


Figure 7.1. Relative quantities of ICETh1 (A) and ICETh2 (B) forms in normal and stress conditions. *attL* detects ICEThs integrated in the chromosome, *attB* detects "scar" produced in the chromosome by recombination after ICE excision and *attI* detects element excised. Black bars show control cultures and grey bars show results of cultures exposed to UV light. Replicates of control cultures were normalized to 1 on each form. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two) between control and UV treated samples of the same form detection. Samples have been normalized with a PCR product from the alpha subunit of RNA polymerase gene. (n=5)

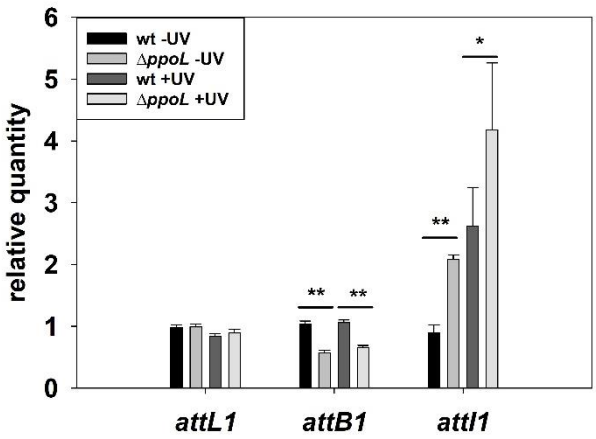
7.1.2 Putative role of PrimPol in replication

These results could suggest some replication of ICEThs under stress conditions. So, different mutants in genes putatively involved in replication were constructed and tested for ICETh1 forms under stress conditions as described above. In these mutants the antibiotic resistance cassette was inserted in the appropriate orientation to allow the transcription of the downstream genes.

First, a knock-out mutant of the PrimPol protein (*TTC0656*) was isolated and subjected to UV treatment and qPCR analysis of ICEThs forms as above. As seen in Figure 7.2A, the amounts of ICETh1 forms seem to be altered in the mutant. The *attI1* form in the non-treated PrimPol mutant exhibits a behavior similar to that of the wild type after UV treatment, with a 2-3 fold increase respect to wild type. This amount of the excised form is again duplicated in the mutant after UV irradiation. At the same time, the *attB1* form is reduced by half in the mutant, with respect to the wild type and does not vary between stress and control conditions. On the other hand, *attL1* behaves exactly like in the wild type.

In the case of ICETh2 (Fig. 7.2B), differences in the excised *attI2* form are even more dramatic. Wild type and mutant control samples differ practically in one log, but surprisingly, the mutant is barely affected by the stress treatment. Behavior of both *attB2* and *attL2* is practically identical than in the wild type.

A



B

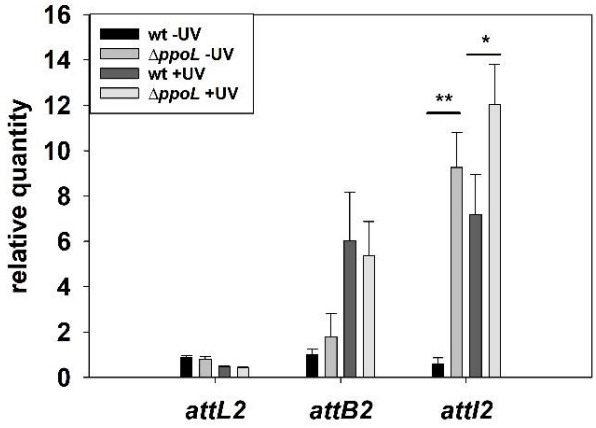


Figure 7.2. Wild type and Δ*ppoL* relative ICETh1 (A) and ICETh2 (B) forms: *attL* detects ICEThs integrated in the chromosome, *attB* detects "scar" produced in the chromosome by recombination after ICE excision and *attI* detects the excised element. Replicas of control cultures have been normalized to 1 on each ICE. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two) between wild type and mutant under the same conditions for the detection of the same ICE form. Samples have been normalized relative to RNA polymerase alpha subunit gene. (n=5)

According to these results, PrimPol protein is not likely involved in replication of the ICEThs. However, the PrimPol mutant seems to stimulate the excised form of both ICEThs even in the untreated cultures so maybe deletion of this gene could be promoting replication through an unknown mechanism.

Given the fact that a mutant in PrimPol seems to increase the excised form of ICEThs, transjugation frequency of ICETh1 (Δ*pilA4,* 043) was analyzed in comparison to a chromosomal marker (Δ*pilA4,* Δ*gdh or* Δ*TTC0313*), in a wild type and in a Δ*ppoL* background. Cm^R wild type cells were used as recipient cells again. Figure 7.3 shows how ICETh1 transfer is not affected by *ppoL* deletion, while chromosomal marker transfer is slightly stimulated (not significant). Therefore, the increase of circularized ICETh1 detected in a Δ*ppoL* mutant does not affect its transfer to a recipient cell.

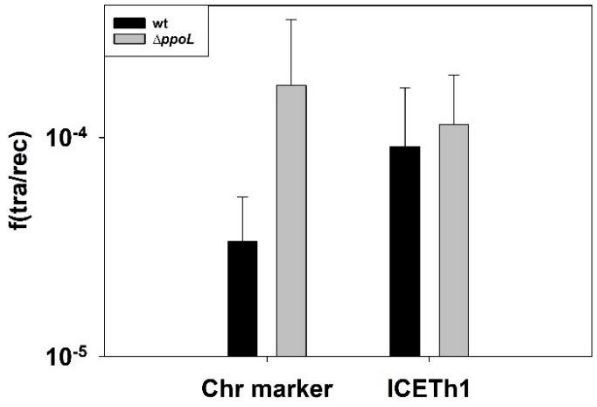


Figure 7.3. Transjugation frequency of a chromosomal marker and ICETh1 to a recipient cell in two different backgrounds, wild type in black and Δ*ppoL* in grey. Transjugation was analyzed as transjugants/recipient cells. In the case of the chromosomal marker transfer, transjugants were selected in Km and Cm in the wt and Hyg and Cm in the mutant. In the case of ICETh1 transfer, both wt and mutant were selected in Hyg and Cm. No significant difference between samples was found.

7.1.3 Analysis of the role of the TOPRIM domain homologue *TTC0657*

Next, a mutant in *TTC0657*, encoding a TOPRIM homologue to domains of the eubacterial primases, was constructed. The same strategy for UV treatment and qPCR described above was applied (Fig. 7.4). In the case of ICETh1 it is noticeable how in the mutant, a 75-fold increase in the "scar" (*attB1*) in comparison to the wild type was detected in both control and UV-treated cultures. It is also noticeable in the mutant, the absence of effects on the amount of excised ICETh1 (*attI1*) detected.

In the case of ICETh2, an increase in the "scar"*attB2* in untreated mutant also occurs but not as dramatic as in ICETh1. In addition, UV treatment just affected slightly the amount of circularized form (*attI2*) level in comparison to the wild type. Finally, the integrated form is also affected in comparison to wild type samples in ICETh2, exhibiting a significant decrease.

These results suggest an implication of the TOPRIM homologue in a hypothetical replication induced by stress of both elements. Actually, it is possible that the ICEThs are being lost in this mutant, as both ICEThs show an increase in the chromosome "scar" (more dramatic in the case of ICETh1), and in the case of ICETh2, in addition, a decrease in the integrated form, suggesting a possible role in ICEThs' maintenance.

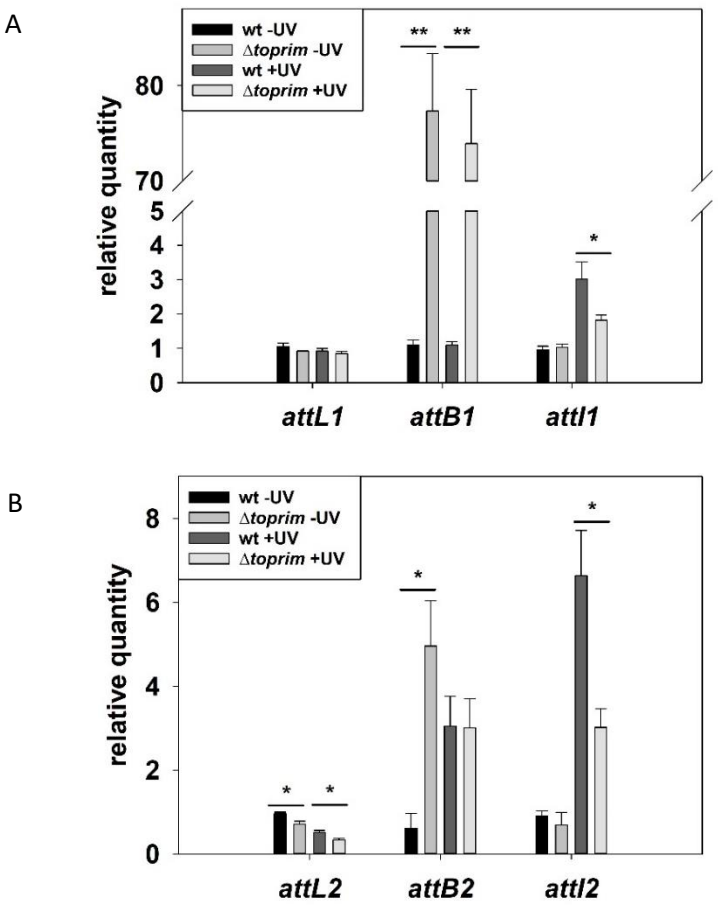


Figure 7.4. Wild type and Δ*toprim* relative ICETh1 (A) and ICETh2 (B) forms quantity: *attL* detects ICEs integrated in the chromosome, *attB* detects the "scar" produced in the chromosome by recombination after ICE excision and *attI* detects the elements once excised. Replicates of control cultures has been normalized to 1 on each form. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two) between the wild type and the mutant under the same conditions and ICE form. Samples have been normalized relative to RNA polymerase alpha subunit, DNA Polymerase III and 16S genes. (n=3)

7.1.4 Role of *TTC0658*

One more mutant was tested. In this case in gene *TTC0658*. This gene encodes an ORF conserved in the *Thermus* genus but without known function, and could be cotranscribed along with *ppoL* and the TOPRIM homologue. The mutant was tested exactly under the same conditions than the other ones. ICETh1 form levels were not significantly affected. *attB1* and *attI1* forms were slightly decreased respect to the wild type cultures in the untreated mutant cultures, but the behavior of the ICEThs forms was similar to that found in the wild type. Something similar was detected for ICETh2. However, the UV treatment in this experiment increased the amount of excised ICETh2,

around 90-fold in the wild type and around 120 times for the mutant. This increase in *attI2* did not affect *attB2* levels while *attL2* levels were slightly decreased (Fig. 7.5).

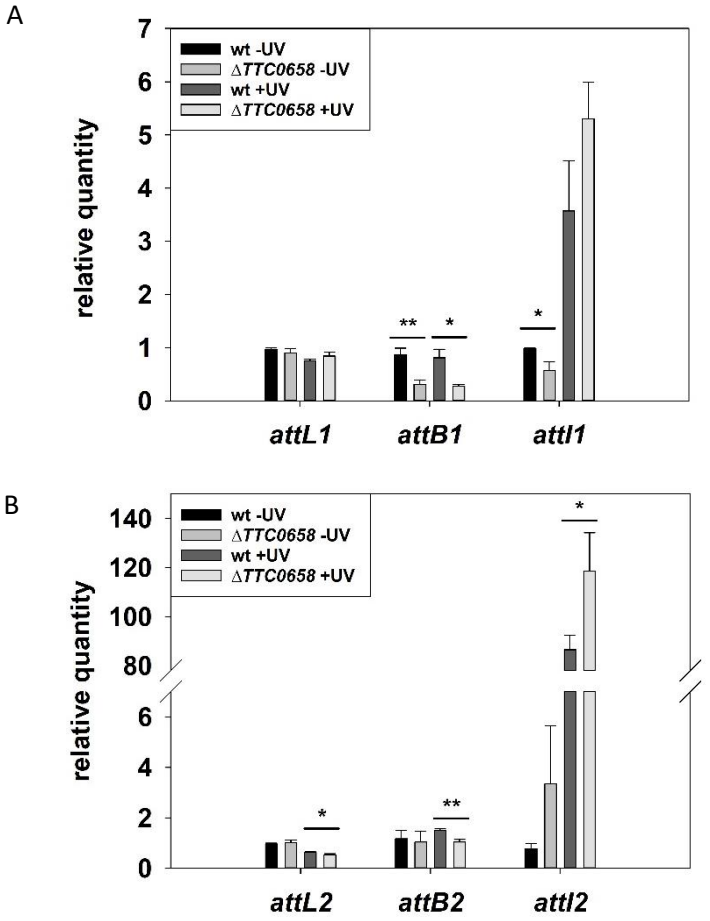


Figure 7.5. Wild type and Δ*TTC0658* relative ICETh1 (A) and ICETh2 (B) forms quantity: *attL* detects ICEThs integrated in the chromosome, *attB* detects "scar" produced in the chromosome by recombination after ICE excision and *attI* detects element excised *per se*. First replica of control cultures has been normalized to 1 on each form. Asterisks indicate significant statistical differences (p-value<0.05 for one and p-value<0.001 for two) between wild type and mutant under the same conditions of the same form detection. Samples have been normalized with polIII and RNA polymerase alpha subunit genes. (n=3)

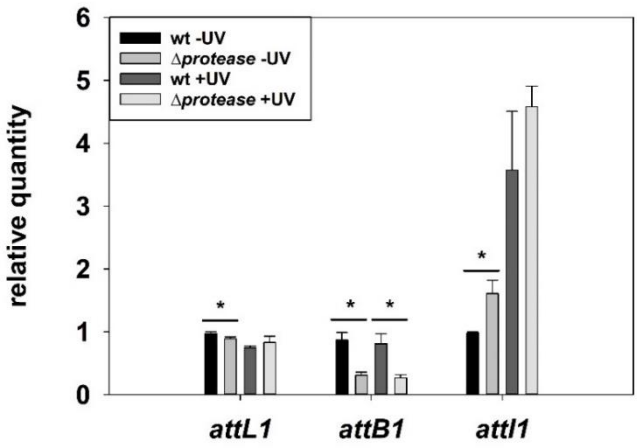
These results do not allow us to conclude that autonomous replication of both ICEThs exists previous to transfer but suggest that it can occur after stress induction. In case of replication, the TOPRIM homologue could be involved in this process, since its absence produce a lower level of both elements.

7.2 Regulation module

ICEs are subjected to signals that often influence their transfer. These signals may activate, repress or influence some modulator genes that in turn can directly or indirectly affect ICE transfer (Beaber and Waldor, 2004; Moon *et al.*, 2005; Bose *et al.*, 2008). A metallopeptidase is encoded in ICETh2, as seen in Chapter 4. This protein could have a regulatory activity as it could be acting like the protease ImmA of ICE*Bs*1 that cleaves the repressor ImmR (which inhibits expression of excisionase *xis* among others) under induction of SOS response (Bose *et al.*, 2008). This repressor-antirepressor system is common in prophages and Gram-positive bacteria. Normally repressor and antirepressor are encoded together. In our case, no putative repressor could be identified within the ICEThs.

The same procedure as in the case of the replication module was applied to a mutant strain in this metallopeptidase (Δ*protease*) and to the wild type strain. As seen in Fig. 7.6, no effect on ICETh1 was observed. The amount of excised ICE in the mutant was higher in both treated and control in comparison to the wild type but the variation was relatively small. A decrease for the "scar" (*attB1*) was also detected in the mutant in comparison to the wt. For ICETh2, only the excised form in the untreated cultures was affected, showing a 25-fold increase. In general, this protein seems not to affect ICETh1 but may have a role in connection to ICETh2 but not as antirepressor, as the effect produced is contrary to the hypothesized.

A



B

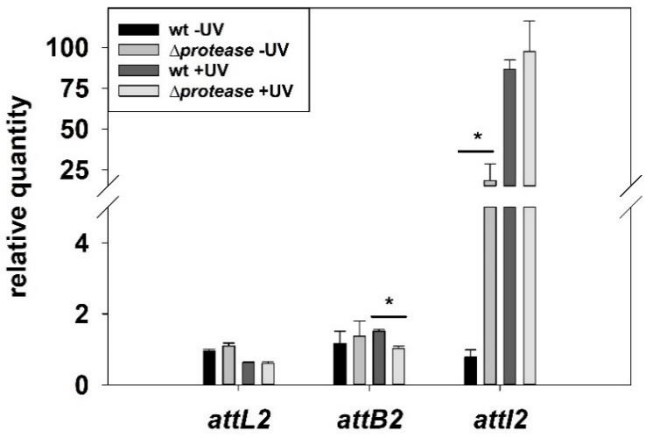


Figure 7.6. Wild type and Δ*protease* relative ICETh1 (A) and ICETh2 (B) forms quantity: *attL* detects ICEThs integrated in the chromosome, *attB* detects "scar" produced in the chromosome by recombination after ICE excision and *attI* detects element excised *per se*. First replica of control cultures has been normalized to 1 on each form. Asterisks indicate significant statistical differences (p-value<0.05) between wt and mutant under the same conditions of the same form detection. Samples have been normalized with RNA polymerase alpha subunit and PolIII genes. (n=3)

7.3 Maintenance module

Finally, a putative maintenance module present in ICETh1 was studied. Such putative module would assure the stable presence of the ICE in which it is encoded and, similarly avoid ICE´s loss after excision.

Alternatively, the presence of a restrictase encoded within ICETh1 could be also promoting the presence of this element in the population as it could serve as a protection against foreign DNA.

A new knockout mutant was constructed for the hypothetical toxin *TTC1885*, whereas the previously generated mutant on the restrictase *TTC1877* was used as well. Mutant and wild type strains were prepared, all of them harboring antibiotic resistance cassettes encoded in ICETh1 (043), and were streaked several times in the presence of the indicated antibiotic in order to ensure that all ICE copies were labelled in this polyploid strain. Then, cultures were grown without selection pressure for 16 generations to analyze ICETh1 loss. To that extent, cultures were diluted when necessary. Four samples were obtained (t0-t3) throughout the process and total DNA was extracted for qPCR quantification of the "scar" in the chromosome (*attB1*). Results of Figure 7.7 show that the three strains seem to have a similar behavior. ICETh1 seems to be lost throughout

the generations at low frequency. In the case of the mutants, the element increases its frequency of loss from t0 to t1 being slightly higher in the case of Δ*tth111II*.

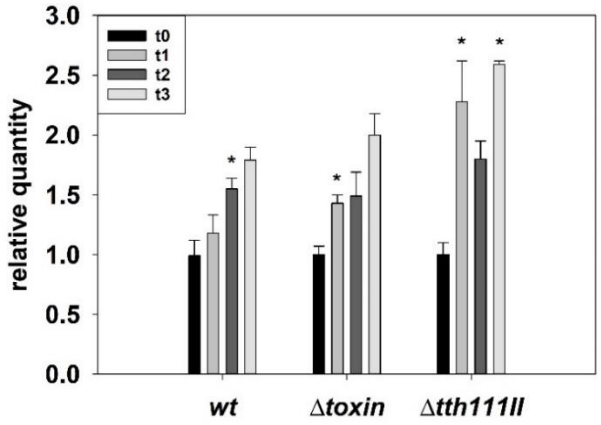


Figure 7.7. ICEth1 loss detection in wild type, Δ*toxin* (*TTC1885*) and Δ*tthIIIII*. Bars show relative quantity of *attB1* detected throughout 16 generations. Each bar of each strain indicates moment of sample extraction, from t0 to t3 from left to right. t0 corresponds to the starting cell culture; t1 was obtained after 4 generations; t2 after 10 and t3 after 16 generations. First replica of each strain at t0 has been normalized to 1. Asterisk indicate significant statistical differences (p-value<0.05) between one bar and the previous one of the same strain. Samples have been normalized with *polIII* and 16S genes. (n=3)

In conclusion, the behavior of both mutants does not suggest any role in maintenance.

As a second approach, we tried to clone the putative toxin in *E. coli* in order to check for its toxicity for the cell. It's been reported before that *T. thermophilus* toxins, despite its activity at high temperatures, can be expressed in a mesophilic bacteria (Fan *et al.*, 2017). For this purpose, the *TTC1885* gene alone and with its putative antitoxin were cloned in parallel in pET28b (+) in the DH5α strain (pIB085 and pIB087 respectively). The resulting plasmids were transformed in the expression strain BL21 and the frequency of transformation was measured (Fig. 7.8). The results show that transformation efficiency was very low when the toxin was cloned alone, but that it increased 10,000-fold when cloned along with the putative antitoxin. Furthermore, toxin expression and induction have been attempted but failed under all the conditions assayed suggesting that it actually has a toxic effect in *E. coli*. These results support that ICETh1 actually encodes a Toxin-Antitoxin module, but with apparently little effects in actual loss of the ICE.

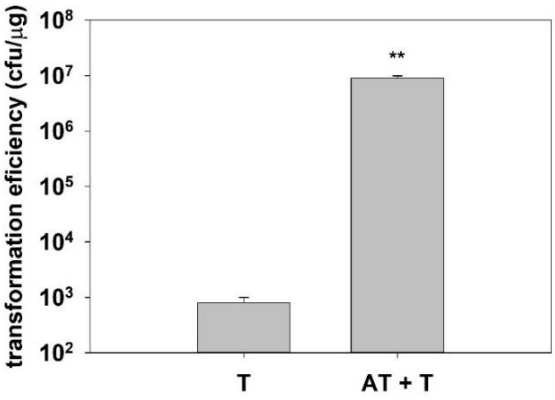


Figure 7.8. Transformation efficiency of *TTC1885* toxin (T) and the same toxin along with its antitoxin (AT + T) in E. coli BL21 strain represented as cfu/μg of DNA. Asterisks indicate significant statistical differences (p-value<0.001) between samples. (n=3)

Chapter 8

Transjugation-generated mosaicity

Chapter 8: Transjugation-generated mosaicity

This chapter has been done in collaboration with Dr. Alba Blesa and Sandra González de la Fuente.

Transjugation has been defined as a horizontal gene transfer process in which the donor cell transfers DNA to a recipient cell, which has to be active and channel this nucleic acid into its cytoplasm through its natural competence machinery (Blesa *et al.*, 2017). In *T. thermophilus* this process apparently transfers DNA from multiple points in the genome simultaneously (Blesa *et al.*, 2015b; Blesa *et al.*, 2017). In this chapter, we analyze by WGS (Whole Genome Sequence) clones derived from transjugation events in order to study whether or not this hypothesis is true.

For this, transjugation assays were performed using *T. thermophilus* HB27 Δ*gdh* (labelled with *kat*) as theoretical donor cells (with ICETh1), and two *T. thermophilus* HB8 strains (without ICETh1), one labelled in the chromosome (Δ*TTHA0672*) and the other labelled in the megaplasmid (Δ*TTHB198*), both with *hyg*, as recipients. Transjugation assays were performed as reported in Chapter 3 (Materials and Methods) and transjugants were selected in plates with Hyg and Km. From all the transjugants obtained, 8 clones were selected for sequencing through the Illumina system. In parallel, the HB27 and HB8 parental strains were sequenced as well. Sequencing was carried out at MicrobesNG (Online Resources) and further bioinformatic analysis was carried out at the bioinformatic genomic service of CBMSO.

As seen in Table 8.1, sequence recruitment to reference genomes of the HB27 and HB8 strains in the GenBank (Benson *et al.*, 2013) revealed that all the transjugants sequenced were essentially derivatives of the HB27 strain, as they exhibited 98% identity to this genome. In contrast, only 87% of identity was obtained between the sequences and the genome of the HB8 strain. These data points out that our theoretically “donor” strain containing the ICETh1 actually was the “recipient” strain in our experiment, at least for the *hyg* gene used as selection marker. Actually, the *hyg* gene was located at the expected sites in the HB27 genome (*TTC0313* and *TTP0146*) in the T1-T8 transjugants by searching for sequences overlapping the 5’ and 3’ of this gene marker.

	Total Length	% align vs HB27-GB	% align vs HB8-GB
HB27	1883063	98.45	87.53
HB8 (Chr)	2088721	79.46	99.44
T1	1911781	98.54	87.90
T2	2005105	98.76	87.89
T3	1818453	98.14	86.49
T4	1820113	97.87	85.39
T5	2006735	98.80	87.22
T6	1914997	98.93	87.09
T7	1883440	98.73	85.96
T8	1823386	98.48	85.45

Table 8.1. Sequence recruitment respect to GenBank (GB). T1-T8 indicates each sequenced transjugant. HB27 and HB8 (Chr) are the parental cells. Total length is indicated in base pair. Chr refers to strain labelled in the chromosome, which was also sequenced.

In order to check for the presence of putative genes from the HB8 strain in the transjugants, chromosome and plasmid sequence alignments were performed between the HB27 and HB8 parental strains in search for differential SNPs to be interpreted as signature of one or the other genomes. Comparison of the obtained sequences with the reference genomes showed that both parental strains (HB27 and HB8) used in the experiment, presented 55 and 111 SNPs respectively in the chromosome, and 7 and 3 SNPs respectively in the pTT27 plasmid respect to the corresponding sequences (chromosome and pTT27 plasmid) deposited in the GenBank (AE017221.1 for HB27 chromosome, AE017222.1 for HB27, NC_006461.1 for HB8 chromosome and NC_006462.1 for HB8 pTT27) (Table 8.2). The continuous growth of these strains under laboratory conditions (domestication) or the errors in the reference genome sequence obtained by older methods could be the cause of these differences.

	Variants vs HB27 GB-chromosome	Variants vs HB27 GB-pTT27	Variants vs HB8 GB-Chromosome	Variants vs HB8 GB-pTT27
HB27p	55	7	11948	1703
HB8p	14853	2173	111	3
T1	156	39	13134	1798
T2	54	297	14335	1707
T3	109	58	10541	1539
T4	136	8	10573	1619
T5	54	12	14504	2023
T6	90	7	12514	1883
T7	158	7	11569	1652
T8	136	6	10761	1600

Table 8.2. SNPs identified between the sequenced strains, transjugants (T) and parentals (p) and the sequences deposited in the GenBank (GB).

In any case, as our purpose was to compare the sequence of the parentals used in the experiment with the transjugants obtained, we set our HB27 parental strain genome as the baseline and ignored the differences with the GenBank.

In this context, a total of 14818 SNPs were identified when the sequence of the HB8 parental chromosome was compared with the chromosome of the HB27 parental strain, whereas 2167 differences were found between the respected megaplasmid-associated genes (Table 8.3). This provides an excellent resolution of parenthood of one SNP every 125 bp with a relatively homogeneous distribution (see Fig. 8.1).

Table 8.3 summarizes the SNPs specific from the HB8 strain found in the chromosome and in the megaplasmid of each transjugant in comparison to the parental HB27 strain. The location of these SNPs for chromosomal and megaplasmid genes in the T1-T8 genomes was represented using Circos (see Chapter 3) (Fig. 8.1). As shown in this figure, T1 showed transfer of megaplasmid genes around the resistance marker (around 4 kbp of DNA transferred) as revealed by the presence of 32 SNPs. Interestingly, this transjugant also shows 114 SNPs scattered along five regions of the chromosome, three

corresponding to the transfer of short DNA fragments (0.5 to 2 Kbp), and two of significant extension (approx. 11 and 12 Kbp, respectively). Transjugant T2 showed extensive transfer of megaplasmid genes from the HB8 strain distributed in four patches detected by the presence of 291 HB8-specific SNPs and also, signals of recombination at 6 genes identified by 12 SNPs scattered along the chromosome. T3 also showed several genes in the megaplasmid transferred from HB8, identified by 51 SNPs at regions, which are far from the selection marker, and two long chromosomal patches of 12 and 6 Kbp revealed by the presence of 65 HB8-specific SNPs, with a few more scattered in the chromosome.

Interestingly, T4 and T5 show little signals of DNA transfer (5 and 1 SNPs respectively) of megaplasmid-associated genes from the HB8 except for the presence of the *hyg* marker. However, whereas T4 shows significant transfer of chromosomal genes from the HB8 strain in four different regions from 1 to 11 Kbp that involve 92 out of 99 SNPs found, T5 shows only 11 isolated SNPs spread along the chromosome.

Analysis of transjugant 6-8 revealed SNPs around the *hyg* selected that revealed the recombination of genes surrounding it upon transfer from HB8. In addition, none of these three strains showed any signal of recombination with genes from the megaplasmid of HB8. T6 chromosome contained 48 SNPs from HB8, most of them (32) around the *hyg* marker; although a 2 Kbp region far from this site was detected as originated from HB8. T7 showed 118 SNPs from HB8 distributed in three regions, one expanding 21 Kbp that corresponds to the marker location and two more of 6 and 8 Kbp far in the chromosome. Additional single SNPs were scattered in the chromosome. Finally, T8 shows the expected transfer of a region around the *hyg* marker and also, the transfer of a 15 Kbp segment far away in the chromosome.

Interestingly, as mentioned previously, despite the variety of transferred locus, the *hyg* marker was always found at the expected site, either in the megaplasmid (*TTP0146*, homologue to *TTHB198*) or in the chromosome (*TTC0313* homologue to *TTHA0672*).

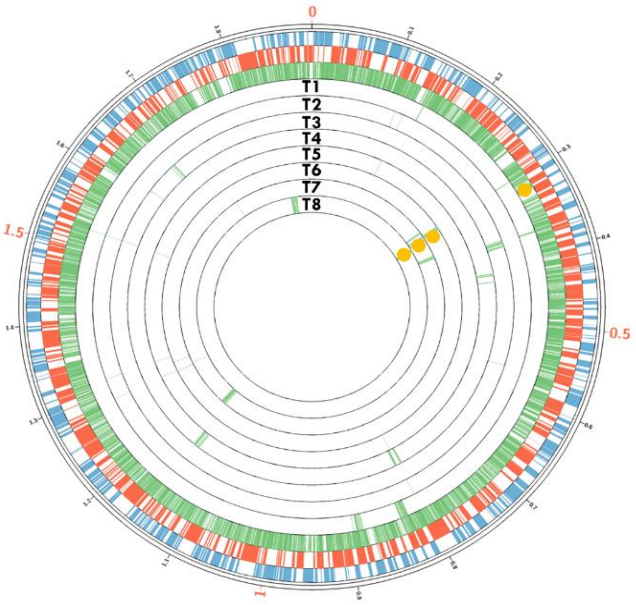
Not only the transfer of genome fragments corresponding to genes with homologues in both strains were detected, but also the transfer of a few entire non-homologous genes specific from HB8 were found in the transjugants. These new acquired genes correspond

to: a putative serine protease (*TTHA0286*) and a hypothetical protein (*TTHA0285*) both transferred in tandem to T3; a putative transcription factor (*TTHB073*) transferred to T2; and two copies of the insertion sequence *ISTh4* transferred to T1. Interestingly, these HB8 genes have been acquired only by transjugants coming from megaplasmid-labelled parental strain, although the low number of sequenced transjugants do not allow us to derive a clear conclusion regarding this trait.

	Chromosome (%)	**pTT27 (%)**	**Non-homologue genes acquired**
HB27	0 (0%)	0 (0%)	-
HB8	14818 (100%)	2167 (100%)	-
T1 (Mpl)	114 (0.77%)	32 (1.47%)	*ISTh4* (2x)
T2 (Mpl)	12 (0.08%)	291 (13.42%)	*TTHB073*
T3 (Mpl)	70 (0.47%)	51 (2.35%)	*TTHA0285* *TTHA0286*
T4 (Mpl)	99 (0.67%)	1 (0.05%)	-
T5 (Mpl)	11 (0.07%)	5 (0.23%)	-
T6 (Chr)	48 (0.32%)	0 (0%)	-
T7 (Chr)	118 (0.8%)	0 (0%)	-
T8 (Chr)	99 (0.67%)	0 (0%)	-

Table 8.3. Variants calling from HB8 respect to the HB27 sequenced parental. T1-T8 indicates individual transjugants; Mpl, strain labelled in the megaplasmid, and Chr, labelled in the chromosome with *hyg*. HB27 and HB8 are the sequenced parental colonies. pTT27 refers to megaplasmid. % indicates the percentage of the genome from recipient HB8 strain present in the analyzed transjugant. Genes acquired, are the genes from HB8 present in the transjugants.

Chromosome



Megaplasmid

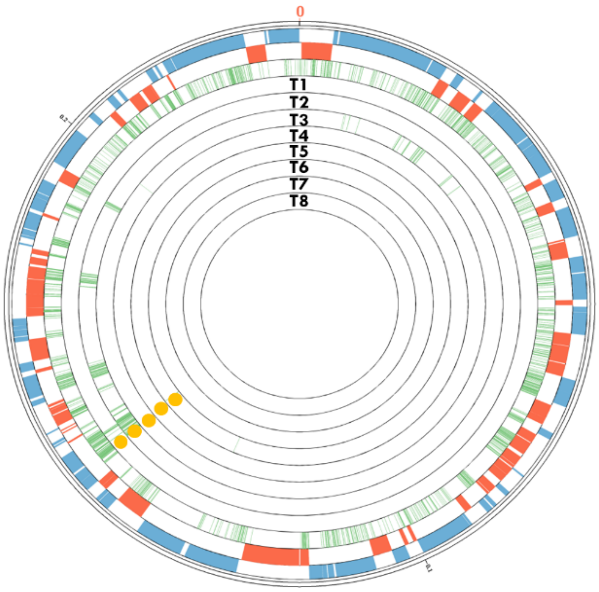


Figure 8.1. HB8 and transjugant variants against HB27 chromosome or pTT27 megaplasmid. Blue and red box shows the HB27 forward and reverse genes respectively. Green lines represent HB8 SNPs variants. First circle showing variants corresponds to HB8. Each transjugant is drawn in different radius within the circle from outside to inside (T1-T8). Yellow circles represent the position of *hyg* resistance cassette.

The results showed above demonstrate that the eight transjugants sequenced were derivatives of the HB27 strain (encoding ICETh1), our theoretical donor, which had

received DNA from the expected recipient (non-encoding ICETh1) by a retro-transfer-like mechanism (discussed Chapter 9).

In order to analyze the unidirectional transfer of genes, a similar experiment was performed using a strain deficient in natural competence (Δ*pilA4*) as a donor. This strain is unable to act as recipient in transjugation, thus avoiding retro-transfer. To that extent, a double HB27 Δ*pilA4*, Δ*gdh* mutant labelled with *kat*, was used as donor and the HB8 Δ*TTHA0672* strain labelled with *hyg* was used as recipient. Five transjugants (T9-T13) were selected and sequenced by Illumina.

Table 8.4 shows the sequence recruitment to the HB8 genome reference. As expected, in this experiment the transjugants selected were derived from the recipient HB8 strain, as more than 99% of their sequences recruited to the reference genome.

	Total Length	% align vs HB27-GB	% align vs HB8-GB
HB8	2088083	79.36	99.70
T9	1978131	82.19	99.54
T10	2047761	84.76	99.53
T11	1872900	82.22	99.51
T12	2060891	83.63	99.52
T13	1984642	81.46	99.19

Table 8.4. Sequence recruitment respect to Gene Bank (GB) reference strains. Numbers are percentage of the sequence that aligns with the indicated reference strains. T9-T13 indicates individual sequenced transjugants. Total length is indicated in base pairs. HB27 is the donor and HB8 is the recipient strain.

As shown before, several SNPs indicative of acquisition of DNA from the donor were detected in the transjugants. The SNPs found, along with genes specific from HB27 also detected in the transjugants, are summarized in Table 8.5. In this case, all the SNPs except for one in T11, were located in the chromosome. In general, all the transjugants exhibited a similar SNP pattern, with the majority of them distributed around the *kat*

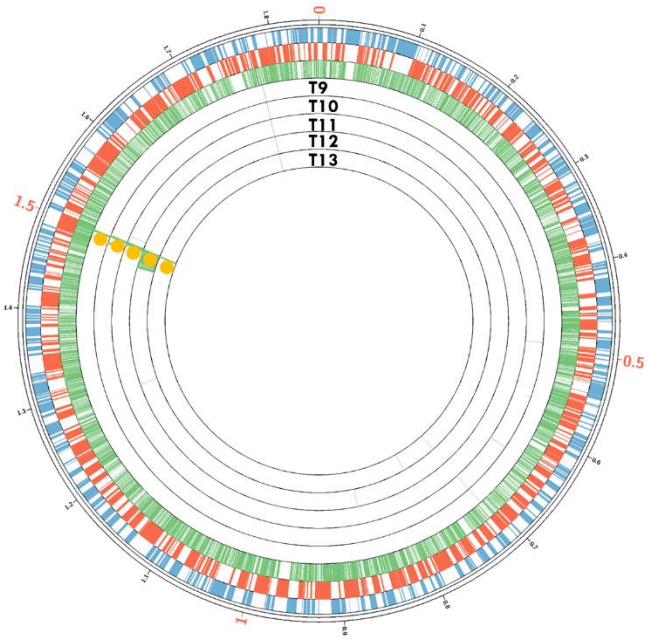
marker, ranging from 2 Kbp as in the case of T9 and T13, 3.2 Kbp in the case of T10, 6.15 Kbp in the case of T11 to 31 Kbp in the case of T12. Furthermore, T10 showed two more regions in which SNPs were grouped, a 3.5 Kbp region close to the *kat* marker region described above, and another small region of 73 bp. Others SNPs detected can be visualized in Figure 8.2.

In this case, the non-homologous genes transferred from HB27 correspond to a putative transporter (*TTC0952*) acquired by T10, a putative transcriptional repressor (*TTC0398*) and the transformation machinery associated protein ComZ (*TTC0857*) to T12, and the gene encoding the TdtA protein (*TTC1879*) in T9. The presence of TdtA, an ICETh1-encoded protein, could suggest the entire integration of ICETh1 in T9, but neither integration site of TdtA nor other parts of the ICE were detected likely due to the low number of reads associated to this gene. This time, no gene coming from the megaplasmid of the HB27 strain could be detected in the megaplasmid of the T9-13 transjugants.

	Chromosome (%)	**pTT27 (%)**	**Non-homologue genes acquired**
HB8	0 (0%)	0 (0%)	-
HB27	14818 (100%)	2167 (100%)	-
T9	20 (0.13%)	0 (0%)	*TTC1879*
T10	64 (0.43%)	0 (0%)	*TTC0952*
T11	42 (0.28%)	1 (0.05%)	-
T12	219 (1.48%)	0 (0%)	*TTC0398* *TTC0857*
T13	20 (0.13%)	0 (0%)	-

Table 8.5. Chromosome or plasmid pTT27 variants coming from HB8 respect to the sequenced parental HB8. T9-T13 indicates individual sequenced transjugants. HB27 is the donor and HB8 is the recipient strain. Genes acquired from HB27 present in the transjugants.

Chromosome



Megaplasmid

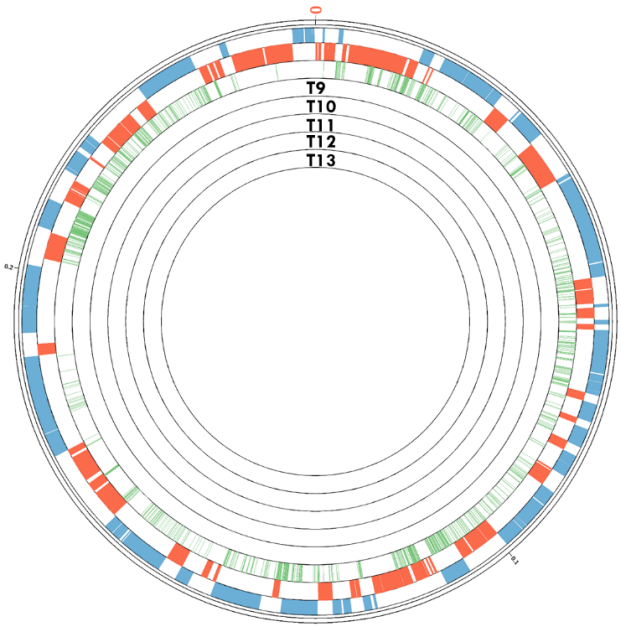


Figure 8.2. HB27 and transjugant variants against HB8 chromosome or pTT27 megaplasmid. Blue and red circles shows the HB8 forward and reverse genes respectively. Green lines represent HB27 variants. First circle showing variants corresponds to HB27. Each transjugant is drawn in different radius within the circle from outside to inside (T9-T13). Yellow circles represent the position of *kat* resistance cassette.

All together, the results show that transjugation produces mosaicism in the transjugants with several DNA fragments recombined into homologue sequences of the recipient genome as detected by SNPs and even new complete genes coming from the donor cell.

Chapter 9

Discussion

Chapter 9: Discussion

Thermus thermophilus strain HB27 encodes two different ICEThs

Searching for Restriction-Modification systems, Furuta *et al.*, 2010 described through *in silico* studies the presence of genes *TTC1877* and *TTC1880* encoding a putative restrictase and a putative methylase, within what seemed an integrative element of *T. thermophilus* HB27, which was absent in the related strain HB8. Later, in a previous work of our group, searching for the protein necessary for the pushing step in transjugation, the DNA translocase TdtA was identified within the same element, that we consequently named Integrative and Conjugative Element 1 (ICETh1).

ICETh1 is the first active ICE described in *T. thermophilus*. It is integrated in the chromosome at a specific *attB* locus and it is able to get excised as a circular form out from the chromosome (Blesa *et al.*, 2017 and Fig. 5.2). Its lower G+C content supports its recent acquisition by HGT. Analysis of its genetic structure revealed that it is formed by different putative modules, all of them likely relevant for its lifecycle (section 4.1). The modules described in this work are, excision/integration, DNA transfer and maintenance. It also encodes a complete and two partial transposases, suggesting that this region may have been the subject to invasion and recombination events after integration, likely mediated by foreign MGEs. Finally, ICETh1 also includes a putative hydrolase.

In the HB27 strain, the presence of a second Integrase (Int2), almost identical to Int1 lead to the finding of ICETh2, an ICE even smaller than ICETh1, which is integrated at a different *attB* site (*attB2*), and with a G+C content similar to that of the species genome. As reported for ICETh1, ICETh2 also excises as an extrachromosomal circular form (Fig. 5.4). In this case, the putative modules encoded are excision/integration, replication and regulation.

Both ICEThs are not conserved across the sequenced *T. thermophilus* strains, but the presence of putative mobile elements seems to be a common feature of the species. For example, ICETh1 is absent in the genome of the HB8 strain, but there seems to be an ICE-like structure, with similarity to ICETh1, integrated at the same *attB1* in the strains

Nar1 and JL-18. On the other hand, there are ICE-like structures integrated in *attB2* in HB8, Nar1 and *SG0*.5JP17-16 strains that also show significant differences respect to ICETh2.

ICETh3 has been identified in this thesis. It is significantly bigger than ICETh1 and ICETh2 and encodes three putative modules that are highly similar to the modules present in ICETh1 and ICETh2: excision/integration, replication and maintenance. As in ICETh2, G+C content of ICETh3 is similar to the chromosomal G+C content but its lifecycle has not been studied.

ICETh1 and ICETh2 have a smaller size than average bacterial ICEs, which range from 18 Kbp to more than 500 Kbp (Johnson and Grossman, 2015). Small ICEs such as ICETh1 and ICETh2 have been found, for example, in the case of *Mesorhizobium ciceri* WSM1271. In this organism, ICEMcSym in its integrated form is dispersed over three segments of 445 Kbp and two small ones of 23 and 7.7 Kbp integrated in different parts of the chromosome (T. L. Haskett *et al.*, 2016). The case of this special ICE will be discussed later. Small ICEs can be found as well in *Actinomycete* (AICEs) as in the case of pMEA300 (13.3 Kbp) of *Amycolatopsis methanolica*, pSE101 (10.9 Kbp) of *Saccharopolyspora erythraea* (Brown, Chiang, *et al.*, 1988; Brown, Tuan, *et al.*, 1988) and the well-studied pSAM2 (10.9 Kbp) of *Streptomyces ambofaciens* (Boccard *et al.*, 1989; Smokvina *et al.*, 1991) in which, at least, an excision/integration, replication and DNA transfer modules are present.

ICETh1 integration/excision module is inactive but ICETh2 module is active for both ICEThs

ICETh1 and ICETh2 are integrated at their specific *attB* sites. Under certain conditions, which are still not precisely determined, they excise from the chromosome and, similarly, they can also get integrated back into the chromosome. For both purposes, ICEs use the machinery encoded in the excision/integration module. The presence of this module in both ICEThs and the fact that their integration sites are different suggested the functionality of both integration modules. However, the absence of an excisionase in ICETh1 has put into doubt this idea. The intracellular mobility assay of

both ICEThs to a plasmid target in different mutant backgrounds (Fig. 6.2) demonstrated that only the module present in ICETh2 was functional, and more surprisingly, that both ICEThs were intracellularly mobilized by the same system. When intracellular mobilization was tested in the HB8 strain with the integration of an artificial mini-ICETh1 (Fig. 6.3), not only Int2 was demonstrated to be capable of mobilizing ICETh1, but also Int3, the tyrosine integrase of ICETh3, revealed its capability to perform the same mobilization.

The apparent lack of activity of Int1 is a puzzling result. The sequence alignment of Int1 and Int2 (Fig. 6.1) reveals 6 amino acid replacements (R1K, R95Q, H190Q, G249R, V253A, K319R) which do not affect the catalytic residues described for tyrosine recombinases (Esposito and Scocca, 1997). Moreover, most of these amino acid changes were present as well in Int3 (R95Q, H190Q, G249R, V253A, K319R), which is active. Hence, a lack of expression of Int1 has been proposed as the cause of its inactivity. However, we have detected transcription at low levels of both *int1* and *int2* with RT-PCR assays (Fig. 6.4), suggesting that the most likely explanation for the lack of activity of Int1 is a defective translation. Indeed, whereas *int2* and its homologue *int3* present identical upstream sequences characterized by the presence of a *Thermus* canonical Shine-Dalgarno sequence (AGGAGG) at an appropriate distance (-11) with respect to the ATG start codon, *int1* exhibits a suboptimal Shine-Dalgarno sequence (GAGGG) at the corresponding position, which suggests a lower translation efficiency under our experimental conditions. However, we cannot discard its putative activity under stress conditions (Figs. 7.1 and 7.2), in which the increase in copy number of circular forms might be due to a higher excision rate (see below) and/or replication.

The fact that the almost identical Exc2 and Exc3 proteins located upstream of *int2* and *int3* genes exhibit a highly conserved helix-turn-helix domain in their C-terminal region (which is representative of excisionases in several phyla) led us to catalog them as excisionases. Our data demonstrates that Exc2 is, actually necessary for ICETh1 and ICETh2 intracellular mobility in *T. thermophilus* HB27 (Fig. 6.2), likely playing a similar role to excisionases present in other ICEs in which they influence the DNA architecture or DNA-protein interactions to promote excision over integration mediated by the integrase (Marra and Scott, 1999; Lewis and Hatfull, 2001). However, it is worth

mentioning that the unusual large distance (120 bp) between the stop codon of *exc* genes and the start codon of their respective *int* genes for *Thermus* spp. suggests that the *int* genes could be expressed from their own promoters, independently of the corresponding excisionases.

Integration of ICEs in more than one *attB* site has been reported in other instances, as the consequence that certain mismatches can be tolerated by the site-specific integrases within the recombination core region. This apparent site promiscuity was detected, in most cases, in mutants lacking the regular *attB* site. This is the case, for example of SXT which integrates at alternative sites in the absence of its primary *attB* site *prfC* (Burrus and Waldor, 2003). Also, ICEclc of *Pseudomonas knackmusii* has demonstrated promiscuity in different $tRNA_{Gly}$ (Sentchilo *et al.*, 2009).

In *T. thermophilus* ICETh1 and ICETh2, a single tyrosine integrase (Int2 or Int3) can recognize different *attI* sites in a kind of "promiscuous" way and recombine them in a precise, site-specific way: each *attI* into its corresponding *attB* target. In Figure 9.1, a comparison of the three *attB* sites to be recognized by Int2 and Int3 is shown. There is a strict conservation of a 7 bp recombination core (ACTTGAA), while SNPs exist in the inverted repeat arms surrounding this recombination core, which likely serve as the recombinase binding sites. Therefore, one possible explanation to the bidirectional specificity of the recombination is that these surrounding sequences would be important for the discrimination of their cognate integration sites. But still, at this point, the possibility that the ICEThs could integrate at alternative *attB* sites in the absence of their primary target site cannot be ruled out. Furthermore, as seen in other recombinases (Biswas *et al.*, 2005), sequences upstream and downstream of the 47-bp *att* sites are needed for the integration reaction, since shorter regions did not lead to integration. For example, in CTnBST from *Bacteroides* spp., the tyrosine integrase IntBST exhibits recombination specificity for different targets as well, and requires 270 bp upstream and downstream of *attBST* for integration (Song *et al.*, 2007). Additionally, the conjugative transposon CTnDOT from *Bacteroides* spp. can target six different *attB* sites that contain the conserved sequence TTTGC at the core (Wood and Gardner, 2015).

Therefore, despite the apparently odd fact that a single integrase, Int2, is able to excise and integrate different ICEThs to their specific *attB* targets, similar observations have

been made in a few mesophilic model systems. This effect is likely due to the ability of the integrases to recognize a core sequence and at the same time display some specificity based on the homology of the surrounding sequences.

Figure 9.1. Alignment of the *attB* sites from the ICETh1 to ICETh3 of *T. thermophilus*. Boxes indicate the inverted regions and the line indicates the recombination core.

ICETh1 and ICETh2 replication could occur prior to transfer

Following the ICEs lifecycle, after excision and prior to transfer, replication by RCR is thought to be a common feature as seen for ICEBs1, Tn916 and ICESt3 among others (Carraro *et al.*, 2011, 2016; Auchtung *et al.*, 2016; Wright and Grossman, 2016). In the specific case of ICEBs1 replication is unidirectional, starts at *oriT* and requires the conjugative relaxase NicK, the ICE-encoded helicase processivity factor HelP, the chromosomally encoded DNA translocase and helicase PcrA, the catalytic subunit of DNA polymerase PolC, and the ß-clamp (DnaN) (Lee *et al.*, 2010; Thomas *et al.*, 2013). The leading strand is presumed to be recircularized by NicK (Khan, 2005). Other ICEs such as ICEclc of *Pseudomonas putida*, have recently demonstrated their capacity to replicate autonomously in transfer-competent cells (Delavat *et al.*, 2019). Additionally, replication has been also reported for ICEs which are transferred in a dsDNA manner as it is the case of pSAM2 in *Streptomyces ambofaciens* (te Poele *et al.*, 2008; Ghinet *et al.*, 2011).

The excised form (*attI*) copy number of both, ICETh1 and ICETh2, suggested no replication for any of them, as their copy number was equal or lower than that of their respective "scars" (*attB*) (Figs. 5.3B and 5.5). However, the induction of a stress response mediated by UV light produced an increase in both *attI1* and *attI2* (Fig. 7.1). This increase cannot be directly attributed to a replication process in ICETh2 given the fact that *attB2* increased proportionally to the increase in *attI2.* However, the increase in *attI1* did not result in higher relative amount of *attB1,* likely, because in normal conditions this "scar"

copy number is one order of magnitude higher than *attI1* (Fig. 5.3B) and so, the addition of the *attB1* copies corresponding to the increase in *attI1* would not be detectable with our methodology. Therefore, this event could be caused either by a replicative increase in the copy number of the excised form or likely by an increase in the excision rate. In other model organisms the induction of stress response induces excision, which normally leads to a higher transfer rate as it's the case of ICESt3 or ICEBs1, among others, when induced by mitomycin C (Auchtung *et al.*, 2005; Carraro *et al.*, 2011). This topic is further discussed below.

In spite of the conclusion outlined above, the presence of a possible replication module, led us to study how these ICEThs behaved in different *T. thermophilus* mutant backgrounds under UV treatment. The first case to study was PrimPol, which is also present in different ICEs from *Actinomycetes* (AICEs) (te Poele *et al.*, 2008), and it has been suggested that, associated with helicases, could participate in a replication initiation complex (Lipps, 2004). In our case, deletion of the *ppoL* gene seems to increase both ICETh1 and ICETh2 excised form. It is interesting to note that, contrary to expected if PrimPol was involved in the ICEThs replication, deletion of *ppoL* increased *attI* copy number respect to wild type even in UV untreated cultures (Fig. 7.2). Actually, the relative decrease in *attB1* copy number and the innocuous effect in *attB2* compared to their respective *attI* in the *ppoL* mutant would suggest the existence of replication for both ICEThs in these conditions. Therefore, our data do not support a role for PrimPol in ICEThs replication.

A possible explanation for the effect of the PrimPol deletion would be based on a possible overexpression of the downstream gene, the TOPRIM homologue (*TTC0657*), due to a polar effect caused by the replacement of *ppoL* gene by the strongly expressed *kat* cassette. This TOPRIM homologue is likely cotranscribed with *ppoL*, as both ORFs overlap. The role of the TOPRIM homologue in replication is supported by two observations: first the increase of the "scar" (*attB*) in both treated and untreated cultures and the decrease in the excised form (*attI*) after stress induction of both ICEThs in the TOPRIM mutant cultures (Fig. 7.4); and second the above discussed increase of the excised forms of the ICEThs in the *ppoL* mutant caused hypothetically by the overexpression of the TOPRIM homologue (Fig. 7.2). Further works leading to the use of

markerless mutants will be needed to check this polar effect hypothesis. Noteworthy, TOPRIM-domain homologues have been rarely found in ICEs, being described only in few cases, for example in ICERanRCAD0133-1 from *Riemerella anatipestifer* (Zhu *et al.*, 2019) and in a putative ICE (or prophage) from *Pseudomonas aeruginosa* in the plasmid pHS87b (Bi *et al.*, 2016), both with an unknown role attributed for the corresponding protein.

Finally, the following gene of the putative operon, *TTC0658,* did not exhibit similarity with any other protein in the GenBank, but it is likely cotranscribed as it overlaps with the previous ORF (*TTC0657*), it exhibits a good RBS and it is conserved in the HB8 strain along with *TTC0657*. The analysis of a deletion mutant did not favor any role in replication for this protein (Fig. 7.5), as its absence only exhibited a slight effect on ICEThs behavior and in the opposite direction than expected by our starting hypothesis.

Besides these results, it is noteworthy that the experiment with *TTC0658* (Fig. 7.5) and also in the experiment performed to test the role of the protease (*TTC0663*) (Fig. 7.6), the stress response mediated by UV leaded to an exponential increase in the excised form of ICETh2 in both, wild type and mutants in comparison to the "scar" in the chromosome. This result is interesting as suggests replication, at least, for ICETh2. A similar result, but less intense, occurs in the *Δttago* mutant (Fig. 5.6). Further experimentation would be needed in this point.

The DNA donation module of ICETh1 is responsible for DNA transfer in transjugation

T. thermophilus transjugation capacity depends on the DNA translocase TdtA encoded in ICETh1. It is present in the strain HB27 and its mutation abolishes transjugation completely (Blesa *et al.*, 2017). This translocase is also present in the NARI strain, which also exhibits an efficient transjugation ability. However, TdtA is not encoded in the HB8 strain, which has a transjugation capacity 2-3 orders of magnitude lower than the HB27 strain. This residual transjugation capacity of HB8 could be explained by the presence of hypothetical alternative transjugation components different from TdtA. Upstream *tdtA*, a type IIG restrictase gene (*TTC1877*) identical to Tth111II is encoded, which

preferentially nicks the top strand at position N11 downstream from its CAARCA target site, followed by *nurA*, a NurA-like nuclease (*TTC1878*). Downstream *tdtA* a hypothetical DNA methylase (*TTC1880*), is encoded as the fourth and last gene of the operon (Fig. 6.6). In our working hypothesis, the restrictase would be in charge of *oriT* recognition, corresponding to the Tth111II target site, as it has been reported that major transfer frequencies correspond to sites near where these restriction sites concentrate (Blesa *et al.*, 2017). This *oriT* recognition would be blocked in case of methylation of the target in which it would be involved the DNA methylase. Somehow, after *oriT* recognition and nicking, substrate DNA would be converted to ssDNA by the NurA-like nuclease. However, it is not clear whether this step occurs, as TdtA structure, could apparently accommodate dsDNA for the transfer process (Blesa *et al.*, 2017). TdtA would push DNA through the cell wall of the donor (probably with the help of unknown accessory proteins) and, finally, the transformation machinery of the recipient cell would be in charge of DNA uptake, ferrying the transferred DNA through its own cell envelope, as reported by Blesa *et al.*, 2015 (Fig. 9.2). The nicked DNA, after transjugation, could be repaired by homologous recombination-mediated DNA repair assisted by HepA (Blesa *et al.*, 2017).

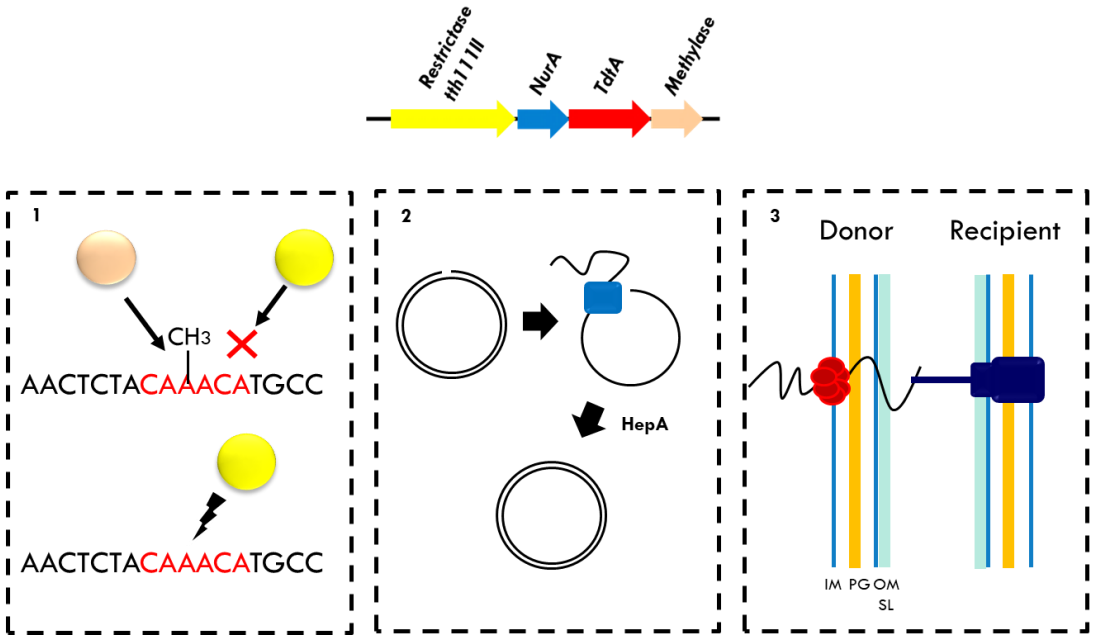


Figure 9.2. Scheme of DNA recognition and transfer during transjugation process. At the top, the four genes encoded on ICETh1 proposed to be involved in transjugation. 1) The restrictase would be in charge of *oriT* recognition producing a nick in on the top strand 11 nucleotides away from its recognition site. Methylase could block *oriT* recognition through methylation of its target sit. 2) NurA hypothetically could transform this dsDNA into ssDNA, as putative substrate for TdtA. Then nicked DNA would be repaired by HepA protein. 3) DNA would be pushed from the donor cell to the recipient cell through TdtA and probably with the help of accessory proteins. Extruded DNA is pulled into the recipient by its own transformation machinery (represented in dark blue). IM: inner membrane; PG: peptidoglycan; OM: outer membrane; SL: S-layer.

Results obtained with mutants in NurA and Tth111II (Fig. 6.7) support their requirement for DNA donation, as both produce a dramatic decrease in transjugation frequencies for chromosomal markers (Blesa *et al.*, 2017). Note that in this experiment, the non-mutant strains are labelled at a chromosomal locus, while the *nurA* and *tth111II* deletions imply labelling of the ICETh1. Since markers in the ICEThs are generally transferred more efficiently that markers in the chromosome, the observed lower transfer rate of the ICE in this experiment means an even more dramatic effect than suggested by direct data comparison.

However, in the present work, even in the absence of Tth111II restriction target, plasmids were transferred to a recipient cell practically with the same efficiency than in the presence of the target. Previous results showed a higher transfer frequency for locus

close to CAARCA sequences (Blesa *et al.*, 2017), therefore there should be alternative ways of DNA recognition for transfer.

Likewise, the mutation of the methylase decreased slightly the transjugation frequency of ICETh1 (Fig. 6.8), which contains 14 Tth111II restriction sites. This again might be suggesting that the hypothesis of restriction site vs *oriT* is not correct.

Alternatively, the topology of DNA could play a major role in the recognition process required for transfer. ICETh1 and ICETh2 exhibited a transfer preference (higher for ICETh1) (Fig. 5.7) with respect to a genomic locus. Interestingly, when ICETh1 excision was abolished by mutation of *int2* or *attR1*, its own transfer was decreased (Fig. 6.5), resulting in a lower transfer frequency than that of a chromosomal marker. This could be related to the fact that a circular DNA can be a better substrate for the transfer machinery. However, it seems in contradiction with this theory the fact that in the *ppoL* mutant, which exhibits a higher copy number of excised ICETh1, the transfer of this element is not increased (Fig. 7.3). Further experiments would be needed to clarify this point.

The hypothetical function of the putative phosphohydrolase (*TTC1884*) in the transfer has not been tested yet but it is tempting to suggest a role in DNA transfer similar to, for example, the case of ICEBs1, in which the phosphohydrolase CwlT is necessary for building an appropriate transfer pore (DeWitt and Grossman, 2014).

Normally ICEs encode the machinery for their own transfer. In our case, the DNA transfer module of ICETh1 can mobilize any locus in the genome (Blesa *et al.*, 2017), besides ICETh2 and itself. This is not a very common feature of ICEs, but there are a few examples of it. For example, SXT (Hochhut *et al.*, 2000), ICEBs1 (Lee *et al.*, 2012) and Tn916 (Naglich and Andrews, 1988) can mobilize plasmids and, additionally, CTnDOT-ERL family of ICEs from *Bacteroides* spp. can replicate genomic islands known as non-replicating *Bacteroides* units (NBUs) (Shoemaker *et al.*, 1986, 1993; Valentine *et al.*, 1988; Stevens *et al.*, 1992). Furthermore, SXT-R91 family of ICEs from *Vibrio* spp. mobilize genomic islands of *Vibrio* and related organisms (Daccord *et al.*, 2010, 2012, 2013). In these two last cases, the genomic islands encoded their own integrase,

necessary for integration in the new host, while excision and transfer depends on the ICE.

ICETh1 and ICETh2 exhibit different maintenance strategies that prevent their loss

Given the fact that replication of ICEThs under normal conditions is very low, if existing at all, there must be some mechanisms to prevent the loss of the element upon excision when the chromosome replicates. Even with loss-preventing systems, ICEs are lost in a small fraction of the population, as reported for SXT/R391 (Wozniak and Waldor, 2009; Carraro *et al.*, 2015). In ICETh1 two putative systems could be involved in the stable maintenance of the ICE in the population: homologues to toxin-antitoxin (TA) and restriction modification (RM) systems (Fig.7.7). According to our data, ICETh1 was lost in certain fraction of the wild type population. However, deletion either of the toxin or the restrictase did not increase its loss rates, maybe due to the low number of generations represented in the experiment. Alternatively, regarding the polyploidy exhibited by *T. thermophilus*, exists the possibility that ICETh1 copy number could be decreasing, but one copy per cell could be maintained. Further experimentation would be needed to corroborate this hypothesis.

Notwithstanding, an *in vivo* test with the TA system demonstrated its activity in *E. coli* (Fig. 7.8). However, despite this result suggests toxicity, it does not ensure the functionality of this system in *T. thermophilus*. This TA system belongs to the *hicAB* type II TA, in which the toxin is a stable translation inhibitor protein, whereas the labile antitoxin is a protein that targets the inhibitor. This mechanism would ensure the post-segregational killing of an ICETh1-free cell upon its loss. TA systems have been reported to be functional in ICEs as in the case of *mosAT* system of SXT/R391 (Wozniak and Waldor, 2009) or the *pezAT* system of the streptococcal Tn5253-like ICE (Mingoia *et al.*, 2014). Interestingly, the crystal structure of the antitoxin of ICETh3, that is 81% identical to that of ICETh1, was obtained in an independent structural study, revealing a novel protein folding, but without showing any clue regarding its actual target to the corresponding toxin (Hattori *et al.*, 2005).

In any case, the presence of a high efficiency transjugation machinery encoded in ICETh1 which can transfer any locus in the genome (Blesa *et al.*, 2017) and therefore can drive the adaptation of the cell population to changes in the environment, plays a major role in ICETh1 loss prevention at bacterial evolutive time scales. Similarly, ICETh2 likely could be evolutionary co-selected with ICETh1 as it improves its spreading mediated by its active excision/integration module (Fig. 6.5).

The TOPRIM homologue (*TTC0657*), found in ICETh2, could also contribute to its stability in the population as its mutation produces a huge loss of ICETh1 as demonstrated by its 80-fold increase in *attB1* (Fig. 7.4) without a concomitant increase in the excised form. This would be suggesting a role for this protein in maintenance of ICETh1 and in a lesser extent of ICETh2.

Besides, partitioning systems prevent the loss of the ICEs by improper distribution among the daughter cells. These systems have been reported for the well studied SXT/R391 which encodes and actin-type *srpMRC* system (Carraro *et al.*, 2015) or the ICE PAPI-1 of *P. aeruginosa* which encodes the putative active partition system Soj, whose deletion leads to high-frequency loss of the ICE (Qiu *et al.*, 2006). There are no homologues of proteins involved in DNA partition encoded in ICEThs. For that reason, ICEThs partition is likely produced in the same way than chromosome segregation (Li, 2019), this is by random segregation.

Regarding PrimPol, this protein has been proposed to act in DNA-DNA interference, possibly by synthesizing DNA guides to load the nuclease Argonaute (A. Blesa, PhD thesis 2016). To test if the phenotype observed in ICEThs could be related to an effect on its DNA-DNA interference mechanism mediated by TtAgo, a mutant without TtAgo was also tested in regard of the copy number of the excised form (Fig. 5.6). However, this mutant only showed an increase in *attI2,* one log over the "scar" (*attB2*). This result would suggest that, on one hand, the ICEThs behavior is barely affected by this type of HGT barrier and, on the other hand, that the effects caused on the ICEThs by the *ttago* and *ppoL* mutants would not be related. Alternatively, the higher excised form of ICETh2 in the mutant could be related to the role of TtAgo as limiter of plasmidic forms.

A regulation mechanism has not been found for ICETh1 and ICETh2

The transmission of ICEs is governed by complex regulatory systems that are activated or repressed by environmental stimuli and can modulate ICE gene expression and transfer. In both ICETh1 and ICETh2, excision and transjugation occurs under normal growth conditions. However, a stress response mediated by UV or the deletion of *ppoL* increases sharply the excised forms of both ICEThs (Figs. 7.1 and 7.2). This increase of the circular forms may occur as a product of either an increase in the excision rate, of a stimulation of replication or both. This stress response might affect to the expression of the integration/excision module or the putative replication module, both encoded in ICETh2. However, as seen in the *ppoL* mutant, transjugation efficiency is not increased by this source of stress (Fig. 7.3), indicating that the increase in DNA excision *per se* is not enough to translate in an increased transfer via regulation of ICE´s transjugation operon. This goes together with the fact that transjugation occurs even when ICETh1 excision is blocked (Fig. 6.5), indicating again that excision is not essential for expression of the DNA transfer module, in contrast to a report by Celli and Trieu-Cuot, 1998 for the conjugative transposon Tn916. In summary, in the ICEs subject of this study, we can conclude that excision is neither necessary nor sufficient for DNA transfer.

As seen in Chapter 1, SetR is the main regulator of SXT/R391 inhibiting excision and transfer, being blocked under SOS response (Beaber and Waldor, 2004). In the case of ICEBs1 the metallopeptidase ImmA cleaves the repressor ImmR leading to excision (Bose *et al.*, 2008). The presence of a metallopeptidase encoded in ICETh2 (*TTC0663*) could suggest the existence of a system similar to that of ICEBs1, despite no putative repressor was identified as hypothetical target. However, the deletion of this metallopeptidase led to an increase of the excised form of ICETh2, indicating that this peptidase could be acting directly or indirectly more like a repressor than as an activator (Fig. 7.6). However, its actual activity or the putative target have not been identified.

Transjugation as a mosaicity generator

The absence of the DNA transfer module encoded in ICETh1, which is present in the HB27 strain and absent in the HB8 strain, may explain the low transfer efficiency of the

HB8 strain as a donor. Furthermore, when ICETh1 is transferred to the HB8 strain, its transjugation efficiency increases dramatically (Blesa *et al.*, 2017). The fact that (Chapter 8) all of the individual genomes sequenced after transjugation between the natural competent HB27 and HB8 strains, are derived from the HB27 strain (Table 8.1) involves a retro-transfer process that could follow the model described in Figure 9.3. In this model, ICETh1 along with other genome sequences (including, theoretically ICETh2) are transferred from the HB27 to the HB8 strain. In this, ICEThs probably could integrate into their *attB* sites and the rest of the genomic fragments would recombine in their homologue sequences in the genome. Then, the HB8 strain becomes a donor due to the presence of ICEth1, and again random genomic DNA fragments (including again ICEThs) can be transferred back in the opposite direction, recombining into the genome of the HB27 strain. This is supported by the fact that the HB8 strain can integrate HB27 DNA as seen in the subsequent experiment with HB8 wild type cells and HB27 Δ*pilA4* (non-recipient) cells (Table 8.4). This model presents some similarities with the *Mycoplasma* model in which MICE is transferred to the MICE-free cell while the later then transfer chromosomal fragments in the opposite direction (Citti *et al.*, 2018). As shown by Blesa *et al.*, 2015, in theory, any locus from the genome could be transferred to the recipient cell.

The transfer process generates a mosaicism in the genome, which likely has a huge impact in the evolution of bacterial genomes, as it occurs in other organisms. This is the case of the mosaic-like genomes of *Mycoplasma* likely produced by retro-transfer that leads to recombinant progeny (Citti *et al.*, 2018). Or the DCT in *Mycobacteria* in which the recombination of the DNA fragments from different sites in the genome leads to a diversity of progeny, each containing a different mosaic-like genome (Gray and Derbyshire, 2018).

Moreover, the retro-transfer detected suggests the involvement of ICETh1 in a gene capture system that could provide its host with higher adaptability to changing environments.

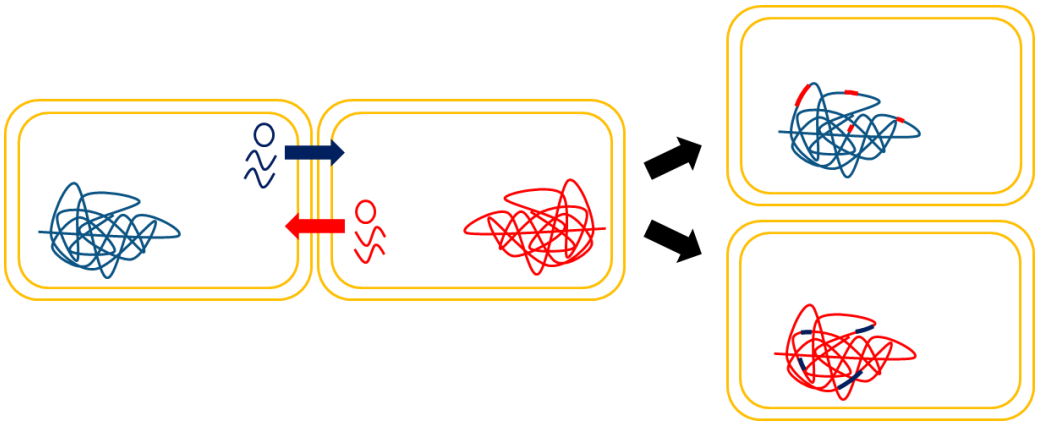


Figure 9.3. Retro-transfer model in *T. thermophilus*. ICEThs along with other DNA fragments are transferred from a donor to a recipient cell in the transjugation process. Once in the recipient cell, they recombine in their homologue sites. As ICETh1 encodes the transjugation machinery, other DNA fragments (including ICEThs) are transferred in the opposite direction and recombine. As a result, the progeny (on the right) exhibits a mosaicism created by the recombination of the DNA fragments coming from different donors.

ICETh2: an IME or part of a bipartite-like ICE with ICETh1?

The case of ICETh2 is especially interesting. As said in Chapter 1, IMEs are integrated in the chromosome, specifically at an *att* site, and encode their own integration/excision module like ICEs, but need the conjugation machinery encoded elsewhere (Bellanger *et al.*, 2014). ICETh2 have demonstrated to fulfill these conditions, as it encodes its own integration/excision module but lacks the DNA transfer module, which is encoded in ICETh1 and necessary for its transfer to a recipient cell. However, ICETh2 or similar ICEThs, as in the case of ICETh3, favors ICETh1 intercellular mobility, providing a functional integration/excision module and possibly, a replication module as well. For that reason, ICETh1 and ICETh2 are proposed to function as a bipartite-like ICE that complement each other to complete their lifecycle, despite the fact that they are transferred separately and not as a single circularized ICE. In the case of ICEMcSym from *M. ciceri* WSM1271, which is the first described tripartite ICE, is composed by three fragments of different lengths integrated at different loci in the chromosome with each fragment flanked by *att* sites corresponding to two different integrases (Fig. 9.4). Therefore, a complex pathway of multiple inversion of chromosomal fragments performed by three different integrases is required to assemble them as a plasmid like-

form. Once in the recipient cell, the ICE disassembles into the three component fragments (T. L. Haskett *et al.*, 2016) to integrate, then, at their corresponding sites. This tripartite ICE seem to be common in *Mesorhizobium* spp. and the sequence of their fragments has a high degree of conservation. The α fragment, which is the largest one, encodes machinery for N_2-fixing symbiosis, for biosynthesis of essential vitamins, excision and transfer (T4SS) but not an integrase (Sullivan *et al.*, 2001, 2002; T. Haskett *et al.*, 2016). However, the β fragment basically carries one of the integrases and the γ fragment encodes the two remaining integrases. This would suggest that the only critical role of the β and γ fragments is to encode the site-specific recombination/excision functions and to maintain the structure of the tripartite ICE (Haskett *et al.*, 2017).

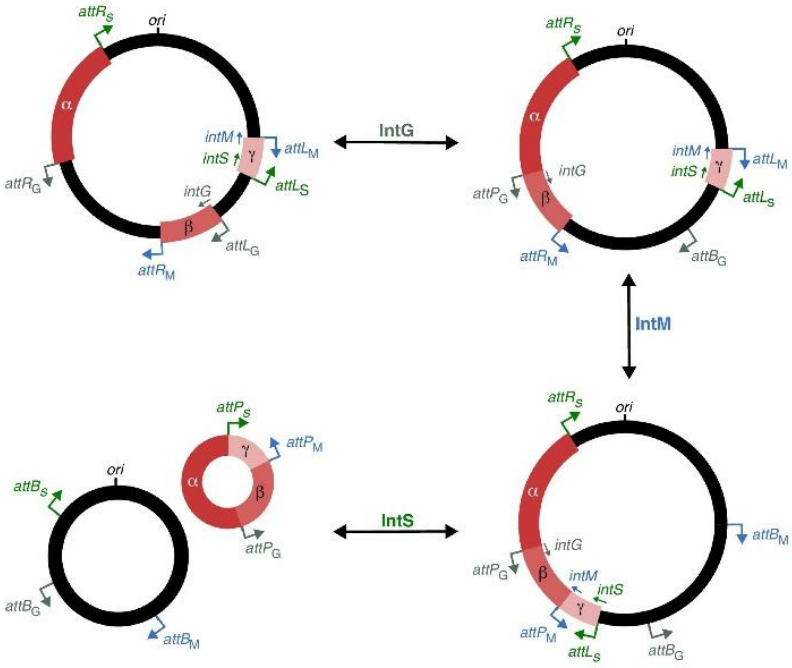


Figure 9.4. Structure of the integrated and excised tripartite ICE of *M. ciceri.* The ICE is fragmented in three segments of different length α, β and γ, which are integrated in different *attB* sites in the chromosome. Upon induction, three integrases mediate the excision of the fragments and through inversion they form a single circularized ICE (Burrus, 2017).

In *T. thermophilus*, as in many thermophiles in general, there is a tendency to reduce the genome as a key factor in the structure-based strategy of adaptation (Berezovsky and Shakhnovich, 2005). Our results suggest the presence of a bipartite-like ICE integrated by two small mobile elements. In this model, ICETh2 encodes the machinery required for the excision of both ICEThs. Prior to transfer, likely under stress

circumstances, replication of the ICEThs could take place using the machinery encoded in ICETh2, at least the TOPRIM homologue. Both ICEThs, along with DNA fragments in the chromosome could be transferred to a recipient cell, this process requiring the DNA transfer module of ICETh1; and once in the recipient cell the two ICEThs could integrate in their respective *attB* sites, again, with the participation of the excision/integration module encoded in ICETh2 or even a host encoded integrase. ICEThs would be maintained in the population by two different mechanisms: i) one is based on the TA system (exclusive for ICETh1) and ii) based on the TOPRIM homologue encoded in ICETh2 (Fig. 9.5).

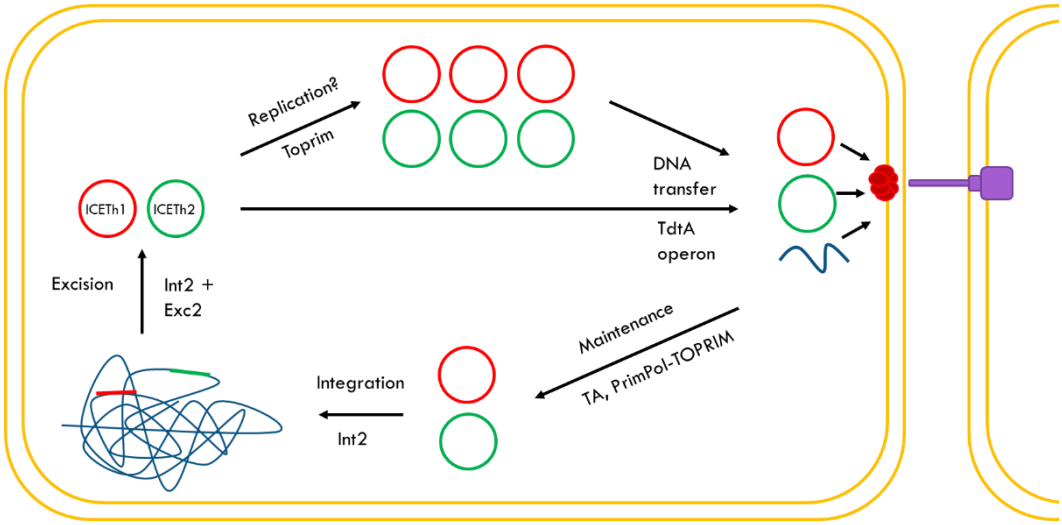


Figure 9.5. Scheme of ICETh1 and ICETh2 lifecycle: Excision of both elements from the chromosome is mediated by the excision/integration module encoded in ICETh2. Replication could occur under certain circumstances and likely, would be mediated by the TOPRIM homologue encoded in ICETh2. Transfer of both ICEThs along with other loci in the genome is performed via the DNA transfer module encoded in ICETh1. Once externalized from the donor cell, the NCA encoded in the genome would uptake the DNA. The Toxin-Antitoxin system (TA) and the TOPRIM homologue could be playing relevant roles in ICEThs maintenance. Finally, both ICEThs can integrate in the corresponding *attB* sites assisted by Int2. Figure elements are not to scale.

Future perspectives

A few questions remain unanswered after this work that require further experimentation.

This is the case of replication, in which the exact conditions in which this event is detected will have to be described. Furthermore, the role of the TOPRIM homologue,

and likely of accessory proteins that participate in this process, will have to be deciphered.

The use of markerless mutants will be necessary to avoid the polar effects, specially in the case of PrimPol. The function of PrimPol is still an enigma, for that reason, future studies have to focus on it.

The DNA transfer process mediated by TdtA will have to be tested as well in order to identify new accessory proteins implied in the process (such as the putative hydrolase) and the specific role of the proteins cotranscribed with TdtA.

In addition, the functionality of the TA module encoded in ICETh1 will have to be tested with different methods in order to know whether it is functional in *T. thermophilus*. Finally, the role of other proteins or systems, such as the TOPRIM homologue, in maintenance of ICEThs will have to be investigated as well.

Conclusions

Conclusions

The main conclusions derived from this work are the following:

1. *Thermus thermophilus* strain HB27 encodes two Integrative and Conjugative Elements (ICEs) coordinated for intracellular and intercellular mobilization, named ICETh1 and ICETh2. We have found also a putative ICE (ICETh3) in the strain HB8.
2. ICETh1 encodes an excision/integration module which is not functional under our experimental conditions and, therefore, requires the corresponding functional module from other ICE to fulfill its lifecycle.
3. ICETh2 provides the functional excision/integration module in the strain HB27 that allows its own intracellular mobilization as well as that of ICETh1, catalyzing the excision and then, the integration of each ICE at their specific corresponding sites.
4. The absence of PrimPol or stress-inducing conditions mediated by UV produce an increase of circular forms of both ICEThs. This can be caused by a higher rate of excision or by the replication of the element, which would be mediated by a TOPRIM-domain homologue encoded in ICETh2.
5. ICETh1 harbors a functional DNA transfer module, in which the transjugation machinery for the pushing step is encoded, that requires the proteins TdtA, NurA and Tth111II. However, the identity of the *oriT* necessary for the transfer process and the role of the methylase cotranscribed with these genes have not yet been elucidated.
6. Maintenance of ICETh1 and, to a lesser extent, of ICETh2 in the population could be mediated by the TOPRIM-domain homologue encoded in ICETh2. Likely the Toxin-Antitoxin system encoded in ICETh1 could have a relevant role in ICETh1 maintenance.
7. Transjugation can occur also as a retro-transfer process, in which ICEThs and other genomic fragments are transferred to a recipient cell. In this recipient cell, the presence of the transjugation machinery allows the transfer of the DNA in the opposite direction; this is, back to the original donor. As a result of this

process, the progeny exhibits mosaicism derived from the integration of the incoming DNA from the recipient.

Conclusiones

Conclusiones

Las principales conclusiones obtenidas en este trabajo son las siguientes:

1. La cepa HB27 de *Thermus thermophilus* codifica dos elementos integrativos conjugativos (ICEs) funcionales y coordinados para la movilización intra e intercelular, denominados ICETh1 e ICETh2. Adicionalmente, otro hipotético elemento (ICETh3) ha sido descubierto en la cepa HB8.
2. ICETh1 codifica un módulo de escisión/integración no funcional en nuestras condiciones experimentales. Por lo tanto, requiere el correspondiente módulo funcional de otro ICE para realizar su ciclo de vida.
3. ICETh2 provee el módulo de escisión/integración funcional en la cepa HB27 para la movilización tanto de ICETh1 como la suya misma, catalizando la escisión y, posteriormente la integración específica de cada ICE en su sitio de integración correspondiente.
4. La ausencia de la proteína PrimPol o las condiciones de estrés producidas por UV producen una mayor tasa de las formas circulares de ambos ICEThs. Esto puede deberse a una mayor tasa de escisión o a la replicación del elemento, en la que estaría implicada una proteína con homología al dominio TOPRIM codificada en ICETh2.
5. ICETh1 alberga un módulo funcional de transferencia de DNA, en el que está codificada la maquinaria de transjugación necesaria para el proceso de extrusión del DNA, que requiere de las proteínas TdtA, NurA y Tth111II. Sin embargo, la identidad de los *oriT* necesarios para el proceso de transferencia y el papel de la metilasa cotranscrita con estos genes todavía no han sido demostrados.
6. El mantenimiento de ICETh1 y, en menor medida, de ICETh2 en la población podría estar mediado por la proteína con homología al dominio TOPRIM codificada en ICETh2. Probablemente, el sistema Toxina-Antitoxina codificado en ICETh1 tenga un papel relevante en el mantenimiento de ICETh1.
7. La transjugación puede ocurrir también como un proceso de retrotransferencia en el cual los ICEThs y otros fragmentos de ADN pueden ser transferidos a una célula receptora. En esta, la presencia de la maquinaria de transjugación permite

la transferencia de ADN en la dirección opuesta, es decir, hacia el donador original. Como consecuencia de este proceso, la progenie exhibe un mosaicismo también en el donador derivado de la integración del ADN entrante desde el receptor.

Bibliography

Bibliography

Aas, F.E., Wolfgang, M., Frye, S., Dunham, S., Løvold, C., and Koomey, M. (2002) Competence for natural transformation in *Neisseria gonorrhoeae*: components of DNA binding and uptake linked to type IV pilus expression. *Mol. Microbiol.* **46**: 749–760.

Acosta, F., Alvarez, L., de Pedro, M.A., and Berenguer, J. (2012) Localized synthesis of the outer envelope from *Thermus thermophilus*. *Extremophiles* **16**: 267–275.

Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., and Shinkai, A. (2010) Transcription Profile of *Thermus thermophilus* CRISPR Systems after Phage Infection. *J. Mol. Biol.* **395**: 270–281.

Alarico, S., Empadinhas, N., Mingote, A., Simões, C., Santos, M.S., and da Costa, M.S. (2007) Mannosylglycerate is essential for osmotic adjustment in *Thermus thermophilus* strains HB27 and RQ-1. *Extremophiles* **11**: 833–840.

Amábile-Cuevas, C.F. and Chicurel, M.E. (1993) Horizontal Gene Transfer. *Am. Sci.* **81**: 332–341.

Andersen, C.L., Jensen, J.L., and Ørntoft, T.F. (2004) Normalization of Real-Time Quantitative Reverse Transcription-PCR Data: A Model-Based Variance Estimation Approach to Identify Genes Suited for Normalization, Applied to Bladder and Colon Cancer Data Sets. *Cancer Res.* **64**: 5245–5250.

Anthony, K.G., Klimke, W.A., Manchak, J., and Frost, L.S. (1999) Comparison of proteins involved in pilus synthesis and mating pair stabilization from the related plasmids F and R100-1: insights into the mechanism of conjugation. *J. Bacteriol.* **181**: 5149–5159.

Aravind, L., Leipe, D.D., and Koonin, E. V (1998) Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res.* **26**: 4205–4213.

Assalkhou, R., Balasingham, S., Collins, R.F., Frye, S.A., Davidsen, T., Benam, A. V, et al. (2007) The outer membrane secretin PilQ from *Neisseria meningitidis* binds DNA. *Microbiology* **153**: 1593–1603.

Auchtung, J.M., Aleksanyan, N., Bulku, A., and Berkmen, M.B. (2016) Biology of ICEBs1, an integrative and conjugative element in *Bacillus subtilis*. *Plasmid* **86**: 14–25.

Auchtung, J.M., Lee, C.A., Monson, R.E., Lehman, A.P., and Grossman, A.D. (2005) Regulation of a *Bacillus subtilis* mobile genetic element by intercellular signaling and the global DNA damage response. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 12554–12559.

Averhoff, B. (2009) Shuffling genes around in hot environments: The unique DNA transporter of *Thermus thermophilus*. *FEMS Microbiol. Rev.* **33**: 611–626.

Balado, M., Lemos, M.L., and Osorio, C.R. (2013) Integrating conjugative elements of the SXT/R391 family from fish-isolated *Vibrios* encode restriction-modification systems that confer resistance to bacteriophages. *FEMS Microbiol. Ecol.* **83**: 457–467.

Baquedano, I., Mencía, M., Blesa, A., Burrus, V., and Berenguer, J. (2019) ICETh1 & ICETh2, two interdependent mobile genetic elements in *Thermus thermophilus* transjugation. Accepted. *Environ. Microbiol.*

Barker, D., Hoff, M., Oliphant, A., and White, R. (1984) A second type II restriction

endonuclease from *Thermus aquaticus* with an unusual sequence specificity. *Nucleic Acids Res.* **12**: 5567–5581.

Bateman, A. (2019) UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**: D506–D515.

Beaber, J.W. and Waldor, M.K. (2004) Identification of operators and promoters that control SXT conjugative transfer. *J. Bacteriol.* **186**: 5945–5949.

Bellanger, X., Morel, C., Gonot, F., Puymege, A., Decaris, B., and Guédon, G. (2011) Site-specific accretion of an integrative conjugative element together with a related genomic island leads to *cis* mobilization and gene capture. *Mol. Microbiol.* **81**: 912–925.

Bellanger, X., Payot, S., Leblond-Bourget, N., and Guédon, G. (2014) Conjugative and mobilizable genomic islands in bacteria: Evolution and diversity. *FEMS Microbiol. Rev.* **38**: 720–760.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Sayers, E.W. (2013) GenBank. *Nucleic Acids Res.* **41**: D36-42.

Berezovsky, I.N. and Shakhnovich, E.I. (2005) Physics and evolution of thermophilic adaptation. *Proc. Natl. Acad. Sci.* **102**: 12742–12747.

Bergé, M., Moscoso, M., Prudhomme, M., Martin, B., and Claverys, J.-P. (2002) Uptake of transforming DNA in Gram-positive bacteria: a view from *Streptococcus pneumoniae*. *Mol. Microbiol.* **45**: 411–421.

Berka, R.M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., et al. (2002) Microarray analysis of the *Bacillus subtilis* K-state: genome-wide expression changes dependent on ComK. *Mol. Microbiol.* **43**: 1331–1345.

Bi, D., Xie, Y., Tai, C., Jiang, X., Zhang, J., Harrison, E.M., et al. (2016) A Site-Specific Integrative Plasmid Found in *Pseudomonas aeruginosa* Clinical Isolate HS87 along with A Plasmid Carrying an Aminoglycoside-Resistant Gene. *PLoS One* **11**: 1–10.

Biswas, T., Aihara, H., Radman-Livaja, M., Filman, D., Landy, A., and Ellenberger, T. (2005) A structural basis for allosteric control of DNA recombination by λ integrase. *Nature* **435**: 1059–1066.

Blesa, A. (2016) Horizontal gene transfer in *Thermus thermophilus*: mechanisms and barriers. PhD thesis. Universidad Autónoma de Madrid.

Blesa, A., Averhoff, B., and Berenguer, J. (2018) Horizontal Gene Transfer in Thermus spp. *Curr. Issues Mol. Biol.* **29**: 23–36.

Blesa, A., Baquedano, I., Quintáns, N.G., Mata, C.P., Castón, J.R., and Berenguer, J. (2017) The transjugation machinery of *Thermus thermophilus*: Identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet.* **13**: 1–22.

Blesa, A. and Berenguer, J. (2019) Alternative Ways to Exchange DNA: Unconventional Conjugation Among Bacteria. In, Villa,T.G. and Viñas,M. (eds), *Horizontal Gene Transfer: Breaking Borders Between Living Kingdoms*. Springer International Publishing, Cham, pp. 77–96.

Blesa, A. and Berenguer, J. (2015) Contribution of vesicle-protected extracellular DNA to horizontal gene transfer in *Thermus* spp. *Int. Microbiol.* **18**: 177–187.

Blesa, A., César, C.E., Averhoff, B., and Berenguer, J. (2015) Noncanonical cell-to-cell DNA

transfer in *Thermus* spp. Is insensitive to argonaute-mediated interference. *J. Bacteriol.* **197**: 138–146.

Blesa, A., Quintans, N.G., Baquedano, I., Mata, C.P., Castón, J.R., and Berenguer, J. (2017) Role of archaeal HerA protein in the biology of the bacterium *Thermus thermophilus*. *Genes (Basel).* **8**:130.

Boccard, F., Smokvina, T., Pernodet, J.L., Friedmann, A., and Guérineau, M. (1989) The integrated conjugative plasmid pSAM2 of *Streptomyces ambofaciens* is related to temperate bacteriophages. *EMBO J.* **8**: 973–980.

Bosch, S., Quesada, J., Esther, S., Lim, J., Berenguer, J., and Hidalgo, A. (2019) Engineered hygromycin phosphotransferase for selection in extreme thermophiles at temperatures above 70^{o}C. Submitted.

Bose, B., Auchtung, J.M., Lee, C.A., and Grossman, A.D. (2008) A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol. Microbiol.* **70**: 570–582.

Brimacombe, C.A., Ding, H., Johnson, J.A., and Beatty, J.T. (2015) Homologues of Genetic Transformation DNA Import Genes Are Required for *Rhodobacter capsulatus* Gene Transfer Agent Recipient Capability Regulated by the Response Regulator CtrA. *J. Bacteriol.* **197**: 2653–2663.

Brochet, M., Couvé, E., Glaser, P., Guédon, G., and Payot, S. (2008) Integrative conjugative elements and related elements are major contributors to the genome diversity of *Streptococcus agalactiae*. *J. Bacteriol.* **190**: 6913–6917.

Brock, T.D. (1978) The Genus *Thermus*. In, *Thermophilic Microorganisms and Life at High Temperatures*. Springer New York, New York, NY, pp. 72–91.

Brown, D.P., Chiang, S.D., Tuan, J.S., and Katz, L. (1988) Site-specific integration in *Saccharopolyspora erythraea* and multisite integration in *Streptomyces lividans* of actinomycete plasmid pSE 101. *J. Bacteriol.* **170**: 2287–2295.

Brown, D.P., Tuan, J.S., Boris, K.A., DeWitt, J.P., Idler, K.B., Chiang, S.J., and Katz, L. (1988) Plasmid-chromosome interactions in *Saccharopolyspora erythraea* and *Streptomyces lividans*. *Dev. Ind. Microbiol.* **29**: 97–105.

Brumm, P.J., Monsma, S., Keough, B., Jasinovica, S., Ferguson, E., Schoenfeld, T., et al. (2015) Complete Genome Sequence of *Thermus aquaticus* Y51MC23. *PLoS One* **10**: 1–30.

Burkhardt, J., Vonck, J., and Averhoff, B. (2011) Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. *J. Biol. Chem.* **286**: 9977–9984.

Burkhardt, J., Vonck, J., Langer, J.D., Salzer, R., and Averhoff, B. (2012) Unusual N-terminal ααβαββα fold of PilQ from *Thermus thermophilus* mediates ring formation and is essential for piliation. *J. Biol. Chem.* **287**: 8484–8494.

Burrus, V. (2017) Mechanisms of stabilization of integrative and conjugative elements. *Curr. Opin. Microbiol.* **38**: 44–50.

Burrus, V. and Waldor, M.K. (2003) Control of SXT Integration and Excision. *J. Bacteriol.* **185**: 5045 LP – 5054.

Burrus, V. and Waldor, M.K. (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res. Microbiol.* **155**: 376–386.

Burton, B. and Dubnau, D. (2010) Membrane-associated DNA transport machines. *Cold Spring Harb. Perspect. Biol.* **2**: a000406.

Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., and Arechaga, I. (2015) Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.* **39**: 81–95.

Carraro, N. and Burrus, V. (2014) Biology of Three ICE Families: SXT/R391, ICEBs1, and ICESt1/ICESt3. *Microbiol. Spectr.* **2**: 101128.

Carraro, N., Libante, V., Morel, C., Charron-Bourgoin, F., Leblond, P., and Guédon, G. (2016) Plasmid-like replication of a minimal streptococcal integrative and conjugative element. *Microbiol. (United Kingdom)* **162**: 622–632.

Carraro, N., Libante, V., Morel, C., Decaris, B., Charron-Bourgoin, F., Leblond, P., and Guédon, G. (2011) Differential regulation of two closely related integrative and conjugative elements from *Streptococcus thermophilus*. *BMC Microbiol.* **11**:238.

Carraro, N., Poulin, D., and Burrus, V. (2015) Replication and active partition of Integrative and conjugative elements (ICEs) of the SXT/R391 family: the line between ICEs and conjugative plasmids is getting thinner. *PLoS Genet.* **11**: e1005298.

Cava, F., Hidalgo, A., and Berenguer, J. (2009) *Thermus thermophilus* as biological model. *Extremophiles* **13**: 213–231.

Cava, F., Pedro, M.A. De, Schwarz, H., Henne, A., and Berenguer, J. (2004) Binding to pyruvylated compounds as an ancestral mechanism to anchor the outer envelope in primitive bacteria. **52**: 677–690.

Cava, F., Zafra, O., Magalon, A., Blasco, F., and Berenguer, J. (2004) A new type of NADH dehydrogenase specific for nitrate respiration in the extreme thermophile *Thermus thermophilus*. *J. Biol. Chem.* **279**: 45369–45378.

Cavalli-Sforza, l. l. (1950) Sexuality of bacteria. In, *Bolletino dell'Istituto sieroterapico milanese*., pp. 281–289.

Celli, J. and Trieu-Cuot, P. (1998) Circularization of Tn916 is required for expression of the transposon-encoded transfer functions: characterization of long tetracycline-inducible transcripts reading through the attachment site. *Mol. Microbiol.* **28**: 103–117.

Chen, I. and Dubnau, D. (2004) DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* **2**: 241–249.

Christie, P.J. (2016) Mosiac Type IV Secretion Systems. *EcoSal Plus* **7**: 1–34.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin).* **6**: 80–92.

Citti, C., Dordet-Frisoni, E., Nouvel, L.X., Kuo, C.H., and Baranowski, E. (2018) Horizontal Gene Transfers in *Mycoplasmas* (*Mollicutes*). *Curr. Issues Mol. Biol.* **29**: 3–22.

Claverys, J.-P., Martin, B., and Polard, P. (2009) The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiol. Rev.* **33**: 643–656.

Claverys, J.-P., Prudhomme, M., and Martin, B. (2006) Induction of Competence Regulons as a General Response to Stress in Gram-Positive Bacteria. *Annu. Rev. Microbiol.* **60**: 451–475.

Collins, R.F., Davidsen, L., Derrick, J.P., Ford, R.C., and Tønjum, T. (2001) Analysis of the PilQ

secretin from *Neisseria meningitidis* by transmission electron microscopy reveals a dodecameric quaternary structure. *J. Bacteriol.* **183**: 3825–3832.

da Costa, M.S., Fernanda Nobre, M., and Rainey, F.A. (2015) *Thermus*. In, *Bergey's Manual of Systematics of Archaea and Bacteria*. American Cancer Society, pp. 1–22.

Curtiss, R. 3rd and Renshaw, J. (1969) F+ strains of *Escherichia coli* K-12 defective in Hfr formation. *Genetics* **63**: 7–26.

Daccord, A., Ceccarelli, D., and Burrus, V. (2010) Integrating conjugative elements of the SXT/R391 family trigger the excision and drive the mobilization of a new class of *Vibrio* genomic islands. *Mol. Microbiol.* **78**: 576–588.

Daccord, A., Ceccarelli, D., Rodrigue, S., and Burrus, V. (2013) Comparative Analysis of Mobilizable Genomic Islands. *J. Bacteriol.* **195**: 606–614.

Daccord, A., Mursell, M., Poulin-Laprade, D., and Burrus, V. (2012) Dynamics of the SetCD-regulated integration and excision of genomic islands mobilized by integrating conjugative elements of the SXT/R391 family. *J. Bacteriol.* **194**: 5794–5802.

Database resources of the National Center for Biotechnology Information. (2016) *Nucleic Acids Res.* **44**: D7-19.

Davies, M.R., Shera, J., Van Domselaar, G.H., Sriprakash, K.S., and McMillan, D.J. (2009) A Novel Integrative Conjugative Element Mediates Genetic Transfer from Group G *Streptococcus* to Other β-Hemolytic Streptococci. *J. Bacteriol.* **191**: 2257 LP – 2265.

Delavat, F., Moritz, R., and van der Meer, J.R. (2019) Transient replication in specialized cells favors transfer of an integrative and conjugative element. *MBio* **10**: 1–15.

Derbyshire, K.M. and Gray, T.A. (2014) Distributive Conjugal Transfer: New Insights into Horizontal Gene Transfer and Genetic Exchange in *Mycobacteria*. *Microbiol. Spectr.* **2**:.

DeWitt, T. and Grossman, A.D. (2014) The bifunctional cell wall hydrolase CwlT is needed for conjugation of the integrative and conjugative element ICEBs1 in *Bacillus subtilis* and B. anthracis. *J. Bacteriol.* **196**: 1588–1596.

Doublet, B., Golding, G.R., Mulvey, M.R., and Cloeckaert, A. (2008) Secondary chromosomal attachment site and tandem integration of the mobilizable *Salmonella* genomic island 1. *PLoS One* **3**: e2060.

Draper, O., César, C.E., Machón, C., de la Cruz, F., and Llosa, M. (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 16385–16390.

Draskovic, I. and Dubnau, D. (2005) Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol. Microbiol.* **55**: 881–896.

Drenkard, E. and Ausubel, F.M. (2002) Pseudomonas biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* **416**: 740–743.

Duffin, P.M. and Barber, D.A. (2016) DprA is required for natural transformation and affects pilin variation in *Neisseria gonorrhoeae*. *Microbiology* **162**: 1620–1628.

Dwivedi, V., Sangwan, N., Nigam, A., Garg, N., Niharika, N., Khurana, P., et al. (2012) Draft Genome Sequence of *Thermus* sp. Strain RL, Isolated from a Hot Water Spring Located atop the Himalayan Ranges at Manikaran, India. *J. Bacteriol.* **194**: 3534 LP – 3534.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., et al. (2019) The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**: D427–D432.

Esposito, D. and Scocca, J.J. (1997) The integrase family of tyrosine recombinases: Evolution of a conserved active site domain. *Nucleic Acids Res.* **25**: 3605–3614.

Fan, Y., Hoshino, T., and Nakamura, A. (2017) Identification of a VapBC toxin–antitoxin system in a thermophilic bacterium *Thermus thermophilus* HB27. *Extremophiles* **21**: 153–161.

Friedrich, A., Hartsch, T., and Averhoff, B. (2001) Natural Transformation in Mesophilic and Thermophilic Bacteria: Identification and Characterization of Novel, Closely Related Competence Genes in *Acinetobacter* sp. Strain BD413 and *Thermus thermophilus* HB27. *Appl. Environ. Microbiol.* **67**: 3140 LP – 3148.

Friedrich, A., Prust, C., Hartsch, T., Henne, A., and Averhoff, B. (2002) Molecular Analyses of the Natural Transformation Machinery and Identification of Pilus Structures in the Extremely Thermophilic Bacterium *Thermus thermophilus* Strain HB27. *Appl. Environ. Microbiol.* **68**: 745 LP – 755.

Friedrich, A., Rumszauer, J., Henne, A., and Averhoff, B. (2003) Pilin-like proteins in the extremely thermophilic bacterium *Thermus thermophilus* HB27: implication in competence for natural transformation and links to type IV pilus biogenesis. *Appl. Environ. Microbiol.* **69**: 3695 LP – 3700.

Fulsundar, S., Harms, K., Flaten, G.E., Johnsen, P.J., Chopade, B.A., and Nielsen, K.M. (2014) Gene Transfer Potential of Outer Membrane Vesicles of *Acinetobacter baylyi* and Effects of Stress on Vesiculation. *Appl. Environ. Microbiol.* **80**: 3469 LP – 3483.

Furste, J.P., Pansegrau, W., Ziegelin, G., Kroger, M., and Lanka, E. (1989) Conjungative transfer of promiscuous IncP plasmids: Interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. U. S. A.* **86**: 1771–1775.

Furuta, Y., Abe, K., and Kobayashi, I. (2010) Genome comparison and context analysis reveals putative mobile forms of restriction-modification systems and related rearrangements. *Nucleic Acids Res.* **38**: 2428–2443.

Garrison, E. and Marth, G. (2012) Haplotype-based variant detection from short-read sequencing. *arXiv e-prints* arXiv:1207.3907.

Ghinet, M.G., Bordeleau, E., Beaudin, J., Brzezinski, R., Roy, S., and Burrus, V. (2011) Uncovering the prevalence and diversity of integrating conjugative elements in *Actinobacteria*. *PLoS One* **6**: e27846.

Goessweiner-Mohr, N., Arends, K., Keller, W., and Grohmann, E. (2013) Conjugative type IV secretion systems in Gram-positive bacteria. *Plasmid* **70**: 289–302.

Gold, V.A., Salzer, R., Averhoff, B., and Kühlbrandt, W. (2015) Structure of a type IV pilus machinery in the open and closed state. *Elife* **4**: 1–12.

De Grado, M., Castán, P., and Berenguer, J. (1999) A high-transformation-efficiency cloning vector for *Thermus thermophilus*. *Plasmid* **42**: 241–245.

Gray, T.A., Clark, R.R., Boucher, N., Lapierre, P., Smith, C., and Derbyshire, K.M. (2016) Intercellular communication and conjugation are mediated by ESX secretion systems in *Mycobacteria*. *Science (80-.).* **354**: 347–350.

Gray, T.A. and Derbyshire, K.M. (2018) Blending genomes: distributive conjugal transfer in *Mycobacteria*, a sexier form of HGT. *Mol. Microbiol.* **108**: 601–613.

Griffith, F. (1928) The Significance of *Pneumococcal* Types. *J. Hyg. (Lond).* **27**: 113–159.

Grohmann, E. (2010) Autonomous plasmid-like replication of *Bacillus* ICEBs1: A general feature of integrative conjugative elements?: MicroCommentary. *Mol. Microbiol.* **75**: 261–263.

Gröschel, M.I., Sayes, F., Simeone, R., Majlessi, L., and Brosch, R. (2016) ESX secretion systems: mycobacterial evolution to counter host immunity. *Nat. Rev. Microbiol.* **14**: 677.

Guédon, G., Libante, V., Coluzzi, C., Payot, S., and Leblond-Bourget, N. (2017) The Obscure World of Integrative and Mobilizable Elements, Highly Widespread Elements that Pirate Bacterial Conjugative Systems. *Genes (Basel).* **8**: 337.

Guglielmini, J., Quintais, L., Garcillán-Barcia, M.P., de la Cruz, F., and Rocha, Eduardo P C (2011) The Repertoire of ICE in Prokaryotes Underscores the Unity, Diversity, and Ubiquity of Conjugation. *PLOS Genet.* **7**: 1–11.

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**: 557–580.

Haskett, T., Wang, P., Ramsay, J., O'Hara, G., Reeve, W., Howieson, J., and Terpolilli, J. (2016) Complete Genome Sequence of *Mesorhizobium ciceri* Strain CC1192, an Efficient Nitrogen-Fixing Microsymbiont of *Cicer arietinum*. *Microbiol. Resour. Announc.* **4**: e00516-16.

Haskett, T.L., Ramsay, J.P., Bekuma, A.A., Sullivan, J.T., O'Hara, G.W., and Terpolilli, J.J. (2017) Evolutionary persistence of tripartite integrative and conjugative elements. *Plasmid* **92**: 30–36.

Haskett, T.L., Terpolilli, J.J., Bekuma, A., O'Hara, G.W., Sullivan, J.T., Wang, P., et al. (2016) Assembly and transfer of tripartite integrative and conjugative genetic elements. *Proc. Natl. Acad. Sci. U. S. A.* **113**: 12268–12273.

Hattori, M., Mizohata, E., Manzoku, M., Bessho, Y., Murayama, K., Terada, T., et al. (2005) Crystal structure of the hypothetical protein TTHA1013 from *Thermus thermophilus* HB8. *Proteins Struct. Funct. Genet.* **61**: 1117–1120.

Hayes, F. (2001) The Horizontal Gene Pool — Bacterial Plasmids and Gene Spread. *Heredity (Edinb).* **86**: 251–252.

He, J., Baldini, R.L., Déziel, E., Saucier, M., Zhang, Q., Liberati, N.T., et al. (2004) The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 2530–2535.

Hidaka, Y., Hasegawa, M., Nakahara, T., and Hoshino, T. (1994) The Entire Population of *Thermus thermophilus* Cells Is Always Competent at Any Growth Phase. *Biosci. Biotechnol. Biochem.* **58**: 1338–1339.

Hochhut, B., Marrero, J., and Waldor, M.K. (2000) Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a constin found in *Vibrio cholerae* O139. *J. Bacteriol.* **182**: 2043–2047.

Humphrey, S.B., Stanton, T.B., Jensen, N.S., and Zuerner, R.L. (1997) Purification and characterization of VSH-1, a generalized transducing bacteriophage of Serpulina hyodysenteriae. *J. Bacteriol.* **179**: 323–329.

Hynes, A.P., Mercer, R.G., Watton, D.E., Buckley, C.B., and Lang, A.S. (2012) DNA packaging bias and differential expression of gene transfer agent genes within a population during

production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Mol. Microbiol.* **85**: 314–325.

Ilangovan, A., Connery, S., and Waksman, G. (2015) Structural biology of the Gram-negative bacterial conjugation systems. *Trends Microbiol.* **23**: 301–310.

Inamine, G.S. and Dubnau, D. (1995) ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *J. Bacteriol.* **177**: 3045–3051.

Inoue, H., Nojima, H., and Okayama, H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene*.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012) A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337**: 816 LP – 821.

Johnson, C.M. and Grossman, A.D. (2015) Integrative and Conjugative Elements (ICEs): What They Do and How They Work. *Annu. Rev. Genet.* **49**: 577–601.

Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.-P. (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* **12**: 181.

Karuppiah, V., Collins, R.F., Thistlethwaite, A., Gao, Y., and Derrick, J.P. (2013) Structure and assembly of an inner membrane platform for initiation of type IV pilus biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**: E4638-47.

Karuppiah, V. and Derrick, J.P. (2011) Structure of the PilM-PilN inner membrane type IV pilus biogenesis complex from *Thermus thermophilus*. *J. Biol. Chem.* **286**: 24434–24442.

Karuppiah, V., Hassan, D., Saleem, M., and Derrick, J.P. (2010) Structure and oligomerization of the PilC type IV pilus biogenesis protein from *Thermus thermophilus*. *Proteins Struct. Funct. Bioinforma.* **78**: 2049–2057.

Khan, S.A. (2005) Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* **53**: 126–136.

Koraimann, G. and Wagner, M.A. (2014) Social behavior and decision making in bacterial conjugation. *Front. Cell. Infect. Microbiol.* **4**: 1–7.

Koyama, Y., Hoshino, T., Tomizuka, N., and Furukawa, K. (1986) Genetic transformation of the extreme thermophile *Thermus thermophilus* and of other Thermus spp. *J. Bacteriol.* **166**: 338–340.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009) Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**: 1639–1645.

Lang, A.S. and Beatty, J.T. (2000) Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. U. S. A.* **97**: 859–864.

Lang, A.S. and Beatty, J.T. (2007) Importance of widespread gene transfer agent genes in *alpha-proteobacteria*. *Trends Microbiol.* **15**: 54–62.

Lang, A.S. and Beatty, J.T. (2001) The gene transfer agent of *Rhodobacter capsulatus* and ``constitutive transduction'' in prokaryotes. *Arch. Microbiol.* **175**: 241–249.

Laurenceau, R., Péhau-Arnaudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., et al.

(2013) A Type IV Pilus Mediates DNA Binding during Natural Transformation in *Streptococcus pneumoniae*. *PLOS Pathog.* **9**: 1–11.

LEDERBERG, J. and TATUM, E.L. (1946) Gene Recombination in *Escherichia coli*. *Nature* **158**: 558.

Lee, C.A., Babic, A., and Grossman, A.D. (2010) Autonomous plasmid-like replication of a conjugative transposon. *Mol. Microbiol.* **75**: 268–279.

Lee, C.A., Thomas, J., and Grossman, A.D. (2012) The *Bacillus subtilis* conjugative transposon ICEBs1 mobilizes plasmids lacking dedicated mobilization functions. *J. Bacteriol.* **194**: 3165–3172.

Lennox, E.S. (1955) Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**: 190–206.

Leong, C.G., Bloomfield, R.A., Boyd, C.A., Dornbusch, A.J., Lieber, L., Liu, F., et al. (2017) The role of core and accessory type IV pilus genes in natural transformation and twitching motility in the bacterium *Acinetobacter baylyi*. *PLoS One* **12**: 1–25.

Lewis, J.A. and Hatfull, G.F. (2001) Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. **29**: 2205–2216.

Li, H. (2019) Random chromosome partitioning in the polyploid bacterium *Thermus thermophilus* HB27. *G3 Genes, Genomes, Genet.* **9**: 1249–1261.

Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Lipps, G. (2004) The replication protein of the *Sulfolobus islandicus* plasmid pRN1. *Biochem. Soc. Trans.* **32**: 240–244.

Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., et al. (2018) ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Res.* **47**: D660–D665.

Livak, K.J. and Schmittgen, T. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCt method. *Methods* **25**: 402–408.

Llamazares, D. (2016) Estudio del efecto de la proteína PrimPol en los procesos de transformación y conjugación de *Thermus thermophilus* HB27. Master degree. Universidad Autónoma de Madrid.

Llosa, M., Gomis-Rüth, F.X., Coll, M., and De la Cruz, F. (2002) Bacterial conjugation: A two-step mechanism for DNA transport. *Mol. Microbiol.* **45**: 1–8.

Londoño-Vallejo, J.A. and Dubnau, D. (1994) Mutation of the putative nucleotide binding site of the *Bacillus subtilis* membrane protein ComFA abolishes the uptake of DNA during transformation. *J. Bacteriol.* **176**: 4642–4645.

Lowe, T.M. and Chan, P.P. (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* **44**: W54–W57.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**: W636—W641.

Marra, D. and Scott, J.R. (1999) Regulation of excision of the conjugative transposon Tn916. **31**: 609–621.

Marrs, B. (1974) Genetic recombination in *Rhodopseudomonas capsulata*. *Proc. Natl. Acad. Sci. U. S. A.* **71**: 971–973.

Mell, J.C., Hall, I.M., and Redfield, R.J. (2012) Defining the DNA uptake specificity of naturally competent *Haemophilus influenzae* cells. *Nucleic Acids Res.* **40**: 8536–8549.

Miller, D.N., Bryant, J.E., Madsen, E.L., and Ghiorse, W.C. (1999) Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl. Environ. Microbiol.* **65**: 4715–4724.

Mingoia, M., Morici, E., Morroni, G., Giovanetti, E., Del Grosso, M., Pantosti, A., and Varaldo, P.E. (2014) Tn5253 family integrative and conjugative elements carrying mef(I) and catQ determinants in *Streptococcus pneumoniae* and *Streptococcus pyogenes*. *Antimicrob. Agents Chemother.* **58**: 5886–5893.

Modi, S.R., Lee, H.H., Spina, C.S., and Collins, J.J. (2013) Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* **499**: 219.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005) Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J. Mol. Evol.* **60**: 174–182.

Moon, K., Shoemaker, N.B., Gardner, J.F., and Salyers, A.A. (2005) Regulation of Excision Genes of the *Bacteroides* Conjugative Transposon CTnDOT. **187**: 5732–5741.

Mortier-Barrière, I., Velten, M., Dupaigne, P., Mirouze, N., Piétrement, O., McGovern, S., et al. (2007) A Key Presynaptic Role in Transformation for a Widespread Bacterial Protein: DprA Conveys Incoming ssDNA to RecA. *Cell* **130**: 824–836.

Naglich, J.G. and Andrews, R.E. (1988) Tn916-dependent conjugal transfer of PC194 and PUB110 from *Bacillus subtilis* into *Bacillus thuringiensis subsp. israelensis. Plasmid* **20**: 113–126.

Navas, L.E., Berretta, M.F., Ortiz, E.M., Benintende, G.B., Amadio, A.F., and Zandomeni, R.O. (2015) Draft Genome Sequence of *Thermus* sp. Isolate 2.9, Obtained from a Hot Water Spring Located in Salta, Argentina. *Genome Announc.* **3**: e01414-14.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304.

Ogura, M., Yamaguchi, H., Kobayashi, K., Ogasawara, N., Fujita, Y., and Tanaka, T. (2002) Whole-genome analysis of genes regulated by the *Bacillus subtilis* competence transcription factor ComK. *J. Bacteriol.* **184**: 2344–2351.

Ohtani, N., Tomita, M., and Itaya, M. (2010) An extreme thermophile, *Thermus thermophilus*, is a polyploid bacterium. *J. Bacteriol.* **192**: 5499–5505.

Omelchenko, M. V, Wolf, Y.I., Gaidamakova, E.K., Matrosova, V.Y., Vasilenko, A., Zhai, M., et al. (2005) Comparative genomics of *Thermus thermophilus* and *Deinococcus radiodurans:* divergent routes of adaptation to thermophily and radiation resistance. *BMC Evol. Biol.* **5**: 57.

Oshima, T. (2007) Unique polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Amino Acids* **33**: 367–372.

Pavlovic, G., Burrus, V., Gintz, B., Decaris, B., and Guédon, G. (2004) Evolution of genomic islands by deletion and tandem accretion by site-specific recombination: ICESt1-related elements from *Streptococcus thermophilus*. *Microbiology* **150**: 759–774.

Penadés, J.R. and Christie, G.E. (2015) The Phage-Inducible Chromosomal Islands: A Family of Highly Evolved Molecular Parasites. *Annu. Rev. Virol.* **2**: 181–201.

Picher, Á.J., Budeus, B., Wafzig, O., Krüger, C., García-Gómez, S., Martínez-Jiménez, M.I., et al. (2016) TruePrime is a novel method for whole-genome amplification from single cells based on TthPrimPol. *Nat. Commun.* **7**: 13296.

te Poele, E.M., Bolhuis, H., and Dijkhuizen, L. (2008) Actinomycete integrative and conjugative elements. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **94**: 127–143.

Qiu, X., Gurkar, A.U., and Lory, S. (2006) Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 19830–19835.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.* **29**: 24–26.

Rose, I., Biuković, G., Aderhold, P., Müller, V., Grüber, G., and Averhoff, B. (2011) Identification and characterization of a unique, zinc-containing transport ATPase essential for natural transformation in *Thermus thermophilus* HB27. *Extremophiles* **15**: 191–202.

Rosenberg, A.H., Lade, B.N., Dao-shan, C., Lin, S.-W., Dunn, J.J., and Studier, F.W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**: 125–135.

Rumszauer, J., Schwarzenlander, C., and Averhoff, B. (2006) Identification, subcellular localization and functional interactions of PilMNOWQ and PilA4 involved in transformation competency and pilus biogenesis in the thermophilic bacterium *Thermus thermophilus* HB27. *FEBS J.* **273**: 3261–3272.

Salzer, R., Herzberg, M., Nies, D.H., Biuković, G., Grüber, G., Müller, V., and Averhoff, B. (2013) The DNA uptake ATPase PilF of *Thermus thermophilus*: a reexamination of the zinc content. *Extremophiles* **17**: 697–698.

Salzer, R., Joos, F., and Averhoff, B. (2014) Type IV pilus biogenesis, twitching motility, and DNA uptake in *Thermus thermophilus*: discrete roles of antagonistic ATPases PilF, PilT1, and PilT2. *Appl. Environ. Microbiol.* **80**: 644–652.

Salzer, R., Kern, T., Joos, F., and Averhoff, B. (2016) The *Thermus thermophilus* comEA/comEC operon is associated with DNA binding and regulation of the DNA translocator and type IV pili. *Environ. Microbiol.* **18**: 65–74.

Sandmann, S., de Graaf, A.O., Karimi, M., van der Reijden, B.A., Hellström-Lindberg, E., Jansen, J.H., and Dugas, M. (2017) Evaluating Variant Calling Tools for Non-Matched Next-Generation Sequencing Data. *Sci. Rep.* **7**: 43169.

Schwarzenlander, C. and Averhoff, B. (2006) Characterization of DNA transport in the thermophilic bacterium *Thermus thermophilus* HB27. *FEBS J.* **273**: 4210–4218.

Schwarzenlander, C., Haase, W., and Averhoff, B. (2009) The role of single subunits of the DNA transport machinery of *Thermus thermophilus* HB27 in DNA binding and transport. *Environ. Microbiol.* **11**: 801–808.

Schwechheimer, C. and Kuehn, M.J. (2015) Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat. Rev. Microbiol.* **13**: 605–619.

Seitz, P., Pezeshgi Modarres, H., Borgeaud, S., Bulushev, R.D., Steinbock, L.J., Radenovic, A., et al. (2014) ComEA Is Essential for the Transfer of External DNA into the Periplasm in Naturally Transformable *Vibrio cholerae* Cells. *PLOS Genet.* **10**: 1–14.

Sentchilo, V., Czechowska, K., Pradervand, N., Minoia, M., Miyazaki, R., and Van Der Meer, J.R. (2009) Intracellular excision and reintegration dynamics of the ICEclc genomic island of *Pseudomonas knackmussii* sp. strain B13. *Mol. Microbiol.* **72**: 1293–1306.

Seth-Smith, H.M.B., Fookes, M.C., Okoro, C.K., Baker, S., Harris, S.R., Scott, P., et al. (2012) Structure, diversity, and mobility of the salmonella pathogenicity island 7 family of integrative and conjugative elements within *enterobacteriaceae*. *J. Bacteriol.* **194**: 1494–1504.

Shoemaker, N.B., Getty, C., Guthrie, E.P., and Salyers, A.A. (1986) Regions in *Bacteroides* plasmids pBFTM10 and pB8-51 that allow *Escherichia coli-Bacteroides* shuttle vectors to be mobilized by IncP plasmids and by a conjugative *Bacteroides* tetracycline resistance element. *J. Bacteriol.* **166**: 959–965.

Shoemaker, N.B., Wang, G.R., Stevens, A.M., and Salyers, A.A. (1993) Excision, transfer, and integration of NBU1, a mobilizable site-selective insertion element. *J. Bacteriol.* **175**: 6578–6587.

Silva, Z., Alarico, S., Nobre, A., Horlacher, R., Marugg, J., Boos, W., et al. (2003) Osmotic Adaptation of *Thermus thermophilus* RQ-1: Lesson from a Mutant Deficient in Synthesis of Trehalose. *J. Bacteriol.* **185**: 5943 LP – 5952.

Smeets, L.C. and Kusters, J.G. (2002) Natural transformation in Helicobacter pilori: DNA transport in an unexpected way. *Trends Microbiol.* **10**: 159–162.

Smillie, C., Garcillan-Barcia, M.P., Francia, M. V., Rocha, E.P.C., and de la Cruz, F. (2010) Mobility of Plasmids. *Microbiol. Mol. Biol. Rev.* **74**: 434–452.

Smokvina, T., Boccard, F., Pernodet, J.-L., Friedmann, A., and Guérineau, M. (1991) Functional analysis of the *Streptomyces ambofaciens* element pSAM2. *Plasmid* **25**: 40–52.

Soler, N., Marguet, E., Verbavatz, J.-M., and Forterre, P. (2008) Virus-like vesicles and extracellular DNA produced by hyperthermophilic archaea of the order *Thermococcales*. *Res. Microbiol.* **159**: 390–399.

Song, B., Shoemaker, N.B., Gardner, J.F., and Salyers, A.A. (2007) Integration site selection by the *Bacteroides* conjugative transposon CTnBST. *J. Bacteriol.* **189**: 6594–6601.

Staals, R.H.J., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., et al. (2014) RNA targeting by the type III-A CRISPR-Cas Csm complex of *Thermus thermophilus*. *Mol. Cell* **56**: 518–530.

Stevens, A.M., Sanders, J.M., Shoemaker, N.B., and Salyers, A.A. (1992) Genes involved in production of plasmidlike forms by a *Bacteroides* conjugal chromosomal element share amino acid homology with two-component regulatory systems. *J. Bacteriol.* **174**: 2935–2942.

Sullivan, J.T., Brown, S.D., Yocum, R.R., and Ronson, C.W. (2001) The bio operon on the acquired symbiosis island of *Mesorhizobium* sp. strain R7A includes a novel gene involved in pimeloyl-CoA synthesis. *Microbiology* **147**: 1315–1322.

Sullivan, J.T., Trzebiatowski, J.R., Cruickshank, R.W., Gouzy, J., Brown, S.D., Elliot, R.M., et al. (2002) Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J. Bacteriol.* **184**: 3086–3095.

Sun, D. (2018) Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. *Front. Microbiol.* **9**: 2154.

Swarts, D.C., Jore, M.M., Westra, E.R., Zhu, Y., Janssen, J.H., Snijders, A.P., et al. (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**: 258–261.

Swarts, D.C., Szczepaniak, M., Sheng, G., Chandradoss, S.D., Zhu, Y., Timmers, E.M., et al. (2017) Autonomous Generation and Loading of DNA Guides by Bacterial Argonaute. *Mol. Cell* **65**: 985-998.e6.

Takeno, M., Taguchi, H., and Akamatsu, T. (2011) Role of ComFA in controlling the DNA uptake rate during transformation of competent *Bacillus subtilis*. *J. Biosci. Bioeng.* **111**: 618–623.

Tatum, E.L. and Lederberg, J. (1947) Gene Recombination in the Bacterium *Escherichia coli*. *J. Bacteriol.* **53**: 673–684.

Teh, B.S., Abdul Rahman, A.Y., Saito, J.A., Hou, S., and Alam, M. (2012) Complete Genome Sequence of the Thermophilic Bacterium *Thermus* sp. Strain CCB_US3_UF1. *J. Bacteriol.* **194**: 1240.

Terui, Y., Ohnuma, M., Hiraga, K., Kawashima, E., and Oshima, T. (2005) Stabilization of nucleic acids by unusual polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Biochem. J.* **388**: 427–433.

Thoma, L. and Muth, G. (2012) Conjugative DNA transfer in *Streptomyces* by TraB: is one protein enough? *FEMS Microbiol. Lett.* **337**: 81–88.

Thoma, L. and Muth, G. (2015) The conjugative DNA-transfer apparatus of *Streptomyces*. *Int. J. Med. Microbiol.* **305**: 224–229.

Thomas, J., Lee, C.A., and Grossman, A.D. (2013) A Conserved Helicase Processivity Factor Is Needed for Conjugation and Replication of an Integrative and Conjugative Element. *PLoS Genet.* **9**: 10.1371.

Tock, M.R. and Dryden, D.T. (2005) The biology of restriction and anti-restriction. *Curr. Opin. Microbiol.* **8**: 466–472.

Turnbull, L., Toyofuku, M., Hynen, A.L., Kurosawa, M., Pessi, G., Petty, N.K., et al. (2016) Explosive cell lysis as a mechanism for the biogenesis of bacterial membrane vesicles and biofilms. *Nat. Commun.* **7**: 11220.

Valentine, P.J., Shoemaker, N.B., and Salyers, A.A. (1988) Mobilization of *Bacteroides* plasmids by *Bacteroides* conjugal elements. *J. Bacteriol.* **170**: 1319–1324.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**: RESEARCH0034.

Veening, J.-W. and Blokesch, M. (2017) Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nat. Rev. Microbiol.* **15**: 621.

Verdú, C., Sanchez, E., Ortega, C., Hidalgo, A., Berenguer, J., and Mencía, M. (2019) A Modular Vector Toolkit with a Tailored Set of Thermosensors To Regulate Gene Expression in *Thermus thermophilus*. *ACS omega* **4**: 14626–14632.

Viana, D., Comos, M., McAdam, P.R., Ward, M.J., Selva, L., Guinane, C.M., et al. (2015) A single natural nucleotide mutation alters bacterial pathogen host tropism. *Nat. Genet.* **47**: 361.

Vieira, J. and Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**: 259–268.

Vogelmann, J., Ammelburg, M., Finger, C., Guezguez, J., Linke, D., Flötenmeyer, M., et al.

(2011) Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. *EMBO J.* **30**: 2246–2254.

Wang, J., Karnati, P.K., Takacs, C.M., Kowalski, J.C., and Derbyshire, K.M. (2005) Chromosomal DNA transfer in *Mycobacterium smegmatis* is mechanistically different from classical Hfr chromosomal DNA transfer. *Mol. Microbiol.* **58**: 280–288.

Wayne, J., Holden, M., and Xu, S. (1997) The Tsp45I restriction–modification system is plasmid-borne within its thermophilic host. *Gene* **202**: 83–88.

Weisburg, W.G., Giovannoni, S.J., and Woese, C.R. (1989) The *Deinococcus-Thermus* Phylum and the Effect of rRNA Composition on Phylogenetic Tree Construction. *Syst. Appl. Microbiol.* **11**: 128–134.

Weldatsadik, R.G., Wang, J., Puhakainen, K., Jiao, H., Jalava, J., Räisänen, K., et al. (2017) Sequence analysis of pooled bacterial samples enables identification of strain variation in group A *Streptococcus*. *Sci. Rep.* **7**: 45771.

Westbye, A.B., Leung, M.M., Florizone, S.M., Taylor, T.A., Johnson, J.A., Fogg, P.C., and Beatty, J.T. (2013) Phosphate concentration and the putative sensor kinase protein CckA modulate cell lysis and release of the *Rhodobacter capsulatus* gene transfer agent. *J. Bacteriol.* **195**: 5025–5040.

Wingett, S.W. and Andrews, S. (2018) FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* **7**: 1338.

Von Wintersdorff, C.J.H., Penders, J., Van Niekerk, J.M., Mills, N.D., Majumder, S., Van Alphen, L.B., et al. (2016) Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Front. Microbiol.* **7**: 1–10.

Wood, M.M. and Gardner, J.F. (2015) The Integration and Excision of CTnDOT. *Mob. DNA III* 183–198.

Wozniak, R.A.F. and Waldor, M.K. (2009) A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. *PLoS Genet.* **5**: e1000439.

Wozniak, Rachel A.F. and Waldor, M.K. (2010) Integrative and conjugative elements:mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* **8**: 552–563.

Wright, L.D. and Grossman, A.D. (2016) Autonomous replication of the conjugative transposon Tn916. *J. Bacteriol.* **198**: 3355–3366.

Wright, L.D., Johnson, C.M., and Grossman, A.D. (2015) Identification of a Single Strand Origin of Replication in the Integrative and Conjugative Element ICEBs1 of *Bacillus subtilis*. *PLOS Genet.* **11**: 1–25.

Yadav, T., Carrasco, B., Hejna, J., Suzuki, Y., Takeyasu, K., and Alonso, J.C. (2013) *Bacillus subtilis* DprA recruits RecA onto single-stranded DNA and mediates annealing of complementary strands coated by SsbB and SsbA. *J. Biol. Chem.* **288**: 22437–22450.

Yadav, T., Carrasco, B., Serrano, E., and Alonso, J.C. (2014) Roles of *Bacillus subtilis* DprA and SsbA in RecA-mediated genetic recombination. *J. Biol. Chem.* **289**: 27640–52.

Yu, M.X., Slater, M.R., and Ackermann, H.-W. (2006) Isolation and characterization of *Thermus* bacteriophages. *Arch. Virol.* **151**: 663–679.

Zafra, O., Ramírez, S., Castán, P., Moreno, R., Cava, F., Vallés, C., et al. (2002) A cytochrome c

encoded by the nar operon is required for the synthesis of active respiratory nitrate reductase in *Thermus thermophilus*. *FEBS Lett.* **523**: 99–102.

Zhu, D., Wan, J., Yang, Z., Xu, J., Wang, M., Jia, R., et al. (2019) First Report of Integrative Conjugative Elements in *Riemerella anatipestifer* Isolates From Ducks in China. *Front. Vet. Sci.* **6**: 128.

Zhu, Z., Guan, S., Robinson, D., Fezzazi, H. El, Quimby, A., and Xu, S.Y. (2014) Characterization of cleavage intermediate and star sites of RM.Tth111II. *Sci. Rep.* **4**: 1–11.

Online resources

Online resources

BacMap	http://bacmap.wishartlab.com/genomes
BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
G+C Content Calculator	https://www.biologicscorp.com/tools/GCContent/
Genome Analysis Toolkit (GATK)	https://software.broadinstitute.org/gatk/
ICEberg	http://ICEberg2/index.php
LAST	http://last.cbrc.jp/
MicrobesNG	https://microbesng.com/
NCBI	https://www.ncbi.nlm.nih.gov/pubmed/
Pfam	https://pfam.xfam.org/
Picard Tools	https://broadinstitute.github.io/picard/
tRNAscan-SE	http://lowelab.ucsc.edu/tRNAscan-SE/
UniProtKB	https://www.uniprot.org/help/uniprotkb
vcffilter	https://github.com/vcflib/vcflib#vcffilter

Annex I

Annex I

Table A1. Oligonucleotides used in this work. Small letters indicate restriction sites and accessory nucleotides for cloning purposes

Primer	Use	Sequence (5'->3')
ATPB Fw	qPCR putative reference gene	GGTTTCCATAGACGAGATCCTG
ATPB Rv	qPCR putative reference gene	AACTCCTCCTCGCTCACGTA
P16S Fw	qPCR putative reference gene	CTCGCAAGCCTTGACAAAAAG
P16S Rv	qPCR putative reference gene	GCAGCAAAAGCCATGCTATCA
P16S Rv	qPCR putative reference gene	GCAGCAAAAGCCATGCTATCA
P306.3	Detection of *attL1* and *attI1* by conventional PCR.	GGTGCTGGACCTGGAAGAAC
P307	Detection of *attR1* and *attI1* by conventional PCR.	CTTGGTCGCCCGAGAGTCAAG
P308	Detection of *attL1* and *attB1* by conventional PCR. Integration detection of pMot Minimal ICE and pMot Minimal ICE YFRQ in HB8	CGTCACCCTGAAGGAGACC
P309	Detection of *attR1* and *attB1* by conventional PCR.	GAGCTTCCGATGGCGTC
P314.2	Transfer module cotranscription detection by PCR.	GCCACACGGTAAGCTCC
P315	Transfer module cotranscription detection by PCR.	GAACACCCAACTTCTGAGCC
P316	Transfer module cotranscription detection by PCR.	GCATTCCCAGATATGAACCG
P317	Transfer module cotranscription detection by PCR.	AACCGTCCAGGAAATAGCG
P318	Transfer module cotranscription detection by PCR.	GGGTGTCGTCAAAGTGGAA
P319	Transfer module cotranscription detection by PCR.	GTCCGTCACATGGCGAAT
P320	Transfer module cotranscription detection by PCR.	ACGGGCAATGCTGCTAC
P321	Transfer module cotranscription detection by PCR.	GCTGAAGCGCACGTAGAAC
P322	Transfer module cotranscription detection by PCR.	GGGAGCGATATGTGCCTTT
P323.2	Transfer module cotranscription detection by PCR.	TTCTGGTGTACGCCCTGG
P324	Construction of pIB009	aactgcagCCGACCCCAATGTGGA
P325	Construction of pIB009	aatctagaGCGGATAGACCGGCTGA

P326	Construction of pIB009	aatctagaGCTTACCCCTCTTTGAGG
P327	Construction of pIB009	aagaattcGGCCTGGCGGCTGAG
P338	Detection of ICETh1 *attB1* in qPCR	GGTTAGAGCGCACGCCTG
P339	Detection of ICETh1 *attB1* in qPCR	GTTGCCGTCCATGCCG
P351	Construction of pIB084	aactgcagGCGTAGGATGTATGAG
P352	Construction of pIB084	aatctagaTTTCATGTCCGCACAC
P353	Construction of pIB084	aatctagaATACTCCGACCTGAGTTG
P354	Construction of pIB084	aagaattcTGGAGGTGGAGAGCGA
P359	Construction of pIB059	aactgcagCTTTCTCTATCCCGT
P359.2	Construction of pIB043	aactgcagGATATGCCCGCTTGC
P360	Construction of pIB059	aatctagaCAGGAGAAGACCAGGG
P361	Construction of Minimal-ICETh1 and pIB059	aatctagaGGCTTTGCTGTCCAT
P362	Construction of pIB039 and pIB059	aagaattcTCGGAGTCCTCCTCG
P372.2	Construction of pIB043	aagaattcGAGCTTCCGATGGCGTC
P373	Construction of pIB043	aatctagaTCTCGGGCGACCAAG
P378	Construction of pIB043	aatctagaCTTGGTCGCCCGAGA
P381	Construction of pIB052	aactgcagGACTCGCTTCCCCTC
P382	Construction of pIB052 and pIB079	aatctagaCACCACCTCCTACGCT
P383.2	Construction of pIB052	aatctagaGGCTAAAGCAATTCCCC
P384	Construction of pIB062 by nested PCR	aagaattcGCCGTTCATGAGCCGC
P384.2	Construction of pIB052 and pIB079	aagaattcGAGGACGAACCAGGAG
P385	Construction of pIB039	aactgcagTTCCTTACAAAAGGC
P386	Construction of pIB039	aatctagaCACGCCCAGAACCC
P387	Construction of pIB039	aatctagaCGCGAGGGGGGAAA
P388	Construction of pIB047	aaactgcagCACAAGCTCCAGC
P389	Construction of pIB047	aaatctagaGTCGATCCGTGGCC
P390	Construction of pIB047	aaatctagaATGAAGCGTACGGAG
P391	Construction of pIB047	aaagaattcCCACTTTGACGACACC
P401	Detection of *attI1* and *attL1* by qPCR	AAGCTTTGTGTTCTGATGCTGGT
P402	Detection of ICETh1 *attL1* in qPCR	TGGCGGGTGCGTGC
P417	Detection of ICETh1 *attI1* in qPCR	CCGCGTCTGGAAGGGATAA
P435	Detection of ICETh2 *attB2* by qPCR	CGAGAAGGAGGTGTGGAAC
P438	Detection of ICETh2 *attB2* by qPCR	TGCCCGGTGTCGTAGA
P439	Construction of pIB060	aactgcagATGGATCGAGCGAGT
P440	Construction of pIB060	aatctagaAAGGTGAGCGTTCGT
P442.2	Construction of pIB060	aagaattcGGGGGAATTGCTTTAG
P443	Construction of pIB061	aactgcagCCAAGATGGAGGAG
P444	Construction of pIB061	aatctagaACGCATGGGGGAGAA
P445	Construction of pIB061	aatctagaTGCGCGAGGGGTTTG
P446	Construction of pIB061 and pIB077	aagaattcCGCTTTTCCAAGAAGG

P451	Detection of ICEThs integration in pIB055 and pIB062	TTCACACAGGAAACAGCTATGAC
P453	Detection of *attL2* and *attB2* by PCR and qPCR	TCGCCGCAATGGAGTTGT
P454	Detection of ICETh2 *attL2* by PCR and qPCR	CCCGAAAATGACCGGCT
P455	Detection of ICETh2 *attR2*	GGGGGAAGCTTTGTGT
P456	Detection of ICETh2 *attR*2 and *attB2*	TCCTGGACCCAGTCG
P461	Construction of pIB055	aatctagaATCGGAAGCGGGGT
P462	Construction of pIB055	aatctagaCGGGGTCGGAGTAC
P466	Construction of Minimal-ICETh1 and pIB078	aatctagaTTTGCCTCCAGGGAC
P467	Construction of pIB060	aatctagaGGTTAGCCGCAAGAC
P472	Construction of pIB062 by nested PCR	aatctagaCGAGAAGGAGGTGTG
P472.2	Construction of pIB062 by nested PCR	CGGGTAGCGTGGTGG
P473	Construction of pIB062 by nested PCR	aatctagaCTTTCCATGAGCTTGG
P474.1	Construction of Minimal-ICETh1	aactgcagGACGCCATGGTGG
P475	Construction of Minimal-ICETh1	aactgcagACTCCTCTACCCC
P476	Construction of Minimal-ICETh1	aactgcagTCAGGCTTGGGCG
P477	Construction of Minimal-ICETh1	aagaattcGGAGCTTCCGATG
P480	Construction of pIB071	aaactagtGTCCACCGCTTTC
P481	Construction of pIB071	aaactagtAGCTATGACATGATTA
P482	Detection of ICETh1 integration in pIB055 and ICETh1 *attI1*	GGGCTTTGCTGTCCATCGTA
P487	Detection of ICETh1 integration in pIB062	CGCACTCTCAAGGTGCAGG
P488	Detection of ICETh2 *attI*2 by PCR and qPCR	AAGCTTTGTGTTCTGATGCTGG
P489	Detection of ICETh2 *attI2* by PCR and qPCR	ACGGGAAGAAAAGGGCTAGTG
P499	Construction of pMH184 hph17 sIFP	aaccatggGCAAAGGAGAAGAA
P500	Construction of pMH184 hph17 sIFP	aagaattcTATTATTTGTAGAGCTCATCC
P501	Construction of pIB070	aactgcagTCATCATAACGCC
P502	Construction of pIB070	aatctagaGAGGTCTTGGAAGG
P503	Construction of pIB070	aatctagaGCCTCTTCAAGGTA
P504	Construction of pIB070	aagaattcGGGCAGACGAACT
P508	Construction of pIB071	aaatgcatTCCCAGGCCACCA
P509	Construction of pIB071	aaccatggCCTCACACCTCC
P514	Construction of Minimal-ICETh1-YFRQ and integration detection of pMot Minimal ICE and pMot Minimal ICE YFRQ in HB8	CTGAATGTGTTCCGCCA

P514	Construction of Minimal-ICETh1-YFRQ (directed mutagenesis) and integration detection of pMot Minimal ICE and pMot Minimal ICE YFRQ in HB8	CTGAATGTGTTCCGCCA
P515	Construction of Minimal-ICETh1-YFRQ	TGGCGGAACACATTCAG
P515	Construction of Minimal-ICETh1-YFRQ (directed mutagenesis)	TGGCGGAACACATTCAG
P516	Construction of Minimal-ICETh1-YFRQ	GCACGACCTTCAACACAC
P516	Construction of Minimal-ICETh1-YFRQ (directed mutagenesis)	GCACGACCTTCAACACAC
P517	Construction of Minimal-ICETh1-YFRQ	GTGTGTTGAAGGTCGTGC
P517	Construction of Minimal-ICETh1-YFRQ (directed mutagenesis)	GTGTGTTGAAGGTCGTGC
P520	Construction of pIB077	aactgcagTCTCGGAAGGCCAA
P521	Construction of pIB077	aatctagaGATGAGCTTCCGCCA
P522	Construction of pIB077	aatctagaGCGTGAGGCGCGTG
P523	Construction of pIB078	aactgcagTCCGAGGTGCCTTC
P524	Construction of pIB078	aatctagaGGGCGGAAGAAGGAAG
P525	Construction of pIB078	aagaattcCTTGGGGTCCTTGAG
P528	Construction of pIB079	aactgcagACCTGGTTCCTGGAG
P529	Construction of pIB079	aatctagaGGCTAAAGCAAATCCCC
P534	Construction of pMotH-SEVA-QX (directed mutagenesis)	CAATGTCGAGCACTTCCG
P535	Construction of pMotH-SEVA-QX (directed mutagenesis)	CGGAAGTGCTCGACATTG
P537	qPCR detection of *exc2*	GCATGCAAGAAGAGGTCCTG
P538	qPCR detection of *exc2*	CACCGTCTTGCGGCTAAC
P539	qPCR detection of *int1*	CCCCTGTTCGATGCTACC
P540	qPCR detection of *int1*	TTGCCGCGCCTCTTA
P542	qPCR detection of *int2*	TTTGGCGCAAGGGAAG
P543	qPCR detection of *int2*	CCTTTGCCGCGTCTTC
P545	qPCR detection of *recA*	AAGGCGATTGAGAAGGAGTTC
P546	qPCR detection of *recA*	CTGCTGCTTGGGCATCTC
P547	qPCR detection of *recO*	CACTCCAGGATGTCCCTCA
P548	qPCR detection of *recO*	CACCACGATGCCCTCTTC
P551	Construction of pIB085	aaccatggCCGACTACGCT
P552	Construction of pIB085	aaaagcttCTAAAACTTTTTTGG
P552.2	Construction of pIB087	aagaattcCTAAAACTTTTTTGG
P553.2	Construction of pIB087	aaccatggTGGGGCTTCAACTA
PPol3 Fw	qPCR putative reference gene	CCGAACGCCTTCTTCTACTC
PPol3 Rv	qPCR putative reference gene	GTGGACGTAGGCTTTGACCT

PRNAPol Fw	qPCR putative reference gene	GCAGCTGGTGATCGAGTTCT
PRNAPol Rv	qPCR putative reference gene	GGGCCTTGAGCTTGGAAT

Annex II

Annex II

Table A2. Buffers and solutions used in this work

Name	Composition	Use
DNA 10X loading buffer	TAE 10X, glycerol 30% (v/v), bromophenol blue 0.25% (w/v), xylene cyanol FF 0.25% (w/v).	DNA electrophoresis: sample preparation
DNase I buffer	Tris-HCl 100 mM pH 7.5, MgCl2 25 mM, CuCl2 1mM	Prevent DNA uptake in conjugation assays
LB medium	Bacto-triptone 10 g/L, yeast extract 5g/L, NaCL 5g/L, pH 7.0	*E. coli* growth medium
PBS 1X	NaCl 150 mM, KCl 25 mM, Na2HPO4 8mM, KH2PO4 1.5 mM	Cell extract resuspension and wash for microscopy preparations
SOB medium	Tryptone 2% (w/v), yeast extract 0.5% (w/v), NaCl 10 mM, KCl 2.5 mM, MgCl2 10 mM, MgSO4 10 mM, MqH2O to 1 L	Broth medium for *E. coli* competent cells
SOC medium	Tryptone 2% (w/v), yeast extract 0.5% (w/v), NaCl 10 mM, KCl 2.5 mM, MgCl2 10 mM, MgSO4 10 mM, glucose 20 mM, MqH2O to 1 L	Transformation of *E. coli* cells
TAE 1X	Tris-acetate 40 mM pH 8; EDTA 1 mM	DNA electrophoresis: gel preparation and running
TB medium	Trypticase 8 g/L, Yeast extract 4 g/L, NaCl 3 g/L in carbonate-rich mineral water, pH 7.5	*T. thermophilus* growth medium
TFBI	RbCl (100 mM), MnCl2·4H2O (50 mM), K Acetate (30 mM), CaCl2·2H2O (10 mM), glycerol (15%) pH 5.8	*E. coli* competence cells preparation buffer
TFBII	MOPS (10 mM), RbCl (10 mM), CaCl2·2H2O (75 mM), glycerol (15%) pH 7	*E. coli* competence cells preparation buffer